EMedium_90_1_2_roberta

August 11, 2022

0.1 Install packages

```
[1]: !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 | 22.59 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
     |                        | 24.2 MB 25.2 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
     |                        | 14.5 MB 34.8 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
     |                        | 689 kB 70.9 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
     |                        | 25.2 MB 1.2 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
     |                        | 1.0 MB 64.8 MB/s
Requirement already satisfied: sklearn==0.0 in
/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
```

```
Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
     |                    | 6.4 MB 51.1 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'…
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 21.50 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 4.9 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 76.0 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
     |                    | 753.4 MB 6.9 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
     |                    | 56 kB 5.6 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
     |                    | 884 kB 67.2 MB/s
Collecting jsonlines==1.2.0
  Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)
```

```
Collecting nltk==3.4.5
  Downloading nltk-3.4.5.zip (1.5 MB)
     |                      | 1.5 MB 61.0 MB/s
Requirement already satisfied: six>=1.5.0 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.15.0)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Collecting blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
     |                      | 3.7 MB 37.3 MB/s
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: 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: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 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)
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 51.6 MB/s
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
     |                      | 184 kB 73.8 MB/s
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: 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: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting tokenizers==0.8.1.rc2
```

```
   Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
      |                          | 3.0 MB 59.1 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting sacremoses
   Downloading sacremoses-0.0.53.tar.gz (880 kB)
      |                          | 880 kB 73.1 MB/s
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: 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: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) … done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449921
sha256=eca1ed445ab09370e8c35ad399d413c9974bb67195eb3b1a380ee6396b0c5b05
  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=5cdde5be4b5782ebdda9f5d7bfd4505020564cda6c91595609f7ad8b7bffcb2e
  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.
```

[1]: True

0.2 Install fp16 optimization

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

```
Cloning into 'apex'…
remote: Enumerating objects: 10117, done.
remote: Counting objects: 100% (233/233), done.
remote: Compressing objects: 100% (165/165), done.
remote: Total 10117 (delta 119), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.97 MiB | 23.47 MiB/s, done.
Resolving deltas: 100% (6927/6927), 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-ur983q3p
Created temporary directory: /tmp/pip-req-tracker-owgsyizj
Initialized build tracking at /tmp/pip-req-tracker-owgsyizj
Created build tracker: /tmp/pip-req-tracker-owgsyizj
Entered build tracker: /tmp/pip-req-tracker-owgsyizj
```

```
Created temporary directory: /tmp/pip-install-yoe3ufrl
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-gvtbv0fv
  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-
owgsyizj'
    Running setup.py (path:/tmp/pip-req-build-gvtbv0fv/setup.py) egg_info for
package from file:///content/ditto/apex
    Created temporary directory: /tmp/pip-pip-egg-info-zt7l8c8q
    Running command python setup.py egg_info


    torch.__version__  = 1.4.0


    running egg_info
    creating /tmp/pip-pip-egg-info-zt7l8c8q/apex.egg-info
    writing /tmp/pip-pip-egg-info-zt7l8c8q/apex.egg-info/PKG-INFO
    writing dependency_links to /tmp/pip-pip-egg-info-zt7l8c8q/apex.egg-
info/dependency_links.txt
    writing top-level names to /tmp/pip-pip-egg-info-zt7l8c8q/apex.egg-
info/top_level.txt
    writing manifest file '/tmp/pip-pip-egg-info-zt7l8c8q/apex.egg-
info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file '/tmp/pip-pip-egg-info-zt7l8c8q/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-gvtbv0fv 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-owgsyizj'
Created temporary directory: /tmp/pip-unpack-yn0pqyyx
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-trbdwvq0
  Building wheel for apex (setup.py) …   Destination directory:
/tmp/pip-wheel-trbdwvq0
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = '"'"'/tmp/pip-req-build-gvtbv0fv/setup.py'"'"';
```

```
__file__='"'"'/tmp/pip-req-build-gvtbv0fv/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-trbdwvq0


  torch.__version__  = 1.4.0


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

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

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

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

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

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

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

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

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

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

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

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

```
  copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
  copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
  copying build/lib/apex/transformer/_data/_batchsampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
  creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
  copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
  copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
  running install_egg_info
  running egg_info
  creating apex.egg-info
  writing apex.egg-info/PKG-INFO
  writing dependency_links to apex.egg-info/dependency_links.txt
  writing top-level names to apex.egg-info/top_level.txt
  writing manifest file 'apex.egg-info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file 'apex.egg-info/SOURCES.txt'
  Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
  running install_scripts
  adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
  creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
  creating '/tmp/pip-wheel-trbdwvq0/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=298244
sha256=e41dfa42233f73920a351de024397b6edf3b0655faeacd92cffceb7232754cb3
    Stored in directory: /tmp/pip-ephem-wheel-cache-
ur983q3p/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-owgsyizj'
/content/ditto
```

[3]:
```python
# some issue with colab
!pip install --upgrade "urllib3==1.25.4" awscli
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
     |                    | 125 kB 4.8 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
     |                    | 3.9 MB 35.0 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
     |                    | 548 kB 73.0 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting botocore==1.27.48
  Downloading botocore-1.27.48-py3-none-any.whl (9.0 MB)
     |                    | 9.0 MB 46.6 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
     |                    | 79 kB 9.6 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
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==1.27.48->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==1.27.48->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
  Attempting uninstall: urllib3
    Found existing installation: urllib3 1.24.3
    Uninstalling urllib3-1.24.3:
      Successfully uninstalled urllib3-1.24.3
  Attempting uninstall: rsa
    Found existing installation: rsa 4.9
    Uninstalling rsa-4.9:
      Successfully uninstalled rsa-4.9
  Attempting uninstall: docutils
    Found existing installation: docutils 0.17.1
    Uninstalling docutils-0.17.1:
      Successfully uninstalled docutils-0.17.1
ERROR: pip's dependency resolver does not currently take into account all

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

dependency conflicts.

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

incompatible.

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

incompatible.
Successfully installed awscli-1.25.48 botocore-1.27.48 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4
```

[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 18 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

[5]:
```
!git clone https://github.com/pauloh48/IC.git
```

```
Cloning into 'IC'…
remote: Enumerating objects: 1014, done.
remote: Counting objects: 100% (84/84), done.
remote: Compressing objects: 100% (74/74), done.
remote: Total 1014 (delta 27), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (1014/1014), 254.12 MiB | 30.79 MiB/s, done.
Resolving deltas: 100% (611/611), done.
Checking out files: 100% (1284/1284), done.
```

0.4 remove e move

[6]:
```
!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
```

[7]:
```
!rm configs.json
!mv -n /content/ditto/IC/datasesErros/EMedium_90_1_2/configs.json /content/
  ↪ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 728kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 27.4MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 34.7MB/s]
Downloading: 100% 501M/501M [00:13<00:00, 37.3MB/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: 0.43594273924827576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216: 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=0.2772277227722772, f1=0.2692307692307693, best_f1=0.2692307692307693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48130878806114197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3010752688172043, f1=0.27083333333333337,
best_f1=0.27083333333333337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4277555048465729
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 3: dev_f1=0.3684210526315789, f1=0.3050847457627119,
best_f1=0.3050847457627119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18049141764640808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37333333333333335, f1=0.32098765432098764,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2842865586280823
/usr/local/lib/python3.7/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=0.2828282828282828, f1=0.2692307692307693,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2859502136707306
/usr/local/lib/python3.7/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=0.2745098039215686, f1=0.2692307692307693,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6642934083938599
/usr/local/lib/python3.7/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=0.2666666666666667, f1=0.2666666666666667,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5031449794769287
/usr/local/lib/python3.7/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=0.2666666666666667, f1=0.2666666666666667,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37663382291793823
/usr/local/lib/python3.7/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=0.2666666666666667, f1=0.2666666666666667,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48185086250305176
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5113305449485779
/usr/local/lib/python3.7/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=0.2828282828282828, f1=0.2692307692307693,
best_f1=0.32098765432098764
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3611753284931183
/usr/local/lib/python3.7/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=0.2828282828282828, f1=0.27184466019417475,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3707781732082367
/usr/local/lib/python3.7/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=0.2745098039215686, f1=0.2692307692307693,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37330326437950134
/usr/local/lib/python3.7/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=0.2745098039215686, f1=0.2692307692307693,
best_f1=0.32098765432098764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44703981280326843
/usr/local/lib/python3.7/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=0.2745098039215686, f1=0.2692307692307693,
best_f1=0.32098765432098764
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, 129912.07it/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 = 0.31746031746031744
real_f1 = 0.35443037974683544
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, 217.89it/s]
```

1.1.2 DBLP-ACM - Running the matcher

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

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

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

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

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

```
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5554894208908081
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4716409146785736
step: 20, loss: 0.5555039048194885
step: 30, loss: 0.2750513255596161
step: 40, loss: 0.16662085056304932
step: 50, loss: 0.054958634078502655
step: 60, loss: 0.07193286716938019
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.08711449056863785
step: 80, loss: 0.04984636977314949
step: 90, loss: 0.1469728648662567
step: 100, loss: 0.16857397556304932
step: 110, loss: 0.08848118036985397
step: 120, loss: 0.0052292910404503345
step: 130, loss: 0.003997360821813345
step: 140, loss: 0.009087640792131424
step: 150, loss: 0.01938074827194214
step: 160, loss: 0.007540115155279636
step: 170, loss: 0.014668596908450127
step: 180, loss: 0.01350807212293148
step: 190, loss: 0.07839395850896835
step: 200, loss: 0.05312453955411911
step: 210, loss: 0.036585886031389236
step: 220, loss: 0.08332120627164841
step: 230, loss: 0.0006560868350788951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9543478260869566, f1=0.9529025191675794,
best_f1=0.9529025191675794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010470468550920486
step: 10, loss: 0.07186511904001236
step: 20, loss: 0.08482853323221207
step: 30, loss: 0.03201451152563095
step: 40, loss: 0.018234627321362495
step: 50, loss: 0.002077008131891489
step: 60, loss: 0.0022111134603619576
step: 70, loss: 0.002538155997171998
```

```
step: 80, loss: 0.003445211797952652
step: 90, loss: 0.011663480661809444
step: 100, loss: 0.0039461879059672356
step: 110, loss: 0.005772701930254698
step: 120, loss: 0.03646800294518471
step: 130, loss: 0.012159515172243118
step: 140, loss: 0.004387620836496353
step: 150, loss: 0.08587641268968582
step: 160, loss: 0.02430099993944168
step: 170, loss: 0.000700528675224632
step: 180, loss: 0.0028855623677372932
step: 190, loss: 0.004568440839648247
step: 200, loss: 0.011473473161458969
step: 210, loss: 0.003964196890592575
step: 220, loss: 0.0007734692189842463
step: 230, loss: 0.0005644452176056802
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9810055865921787, f1=0.9710467706013363,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036329899448901415
step: 10, loss: 0.008639516308903694
step: 20, loss: 0.0011919279349967837
step: 30, loss: 0.000530308170709759
step: 40, loss: 0.028796926140785217
step: 50, loss: 0.027243835851550102
step: 60, loss: 0.0018296679481863976
step: 70, loss: 0.0017446811543777585
step: 80, loss: 0.0006782812997698784
step: 90, loss: 0.0051932926289737225
step: 100, loss: 0.0014320010086521506
step: 110, loss: 0.001884756376966834
step: 120, loss: 0.00044909960706718266
step: 130, loss: 0.001200154540129006
step: 140, loss: 0.0019504404626786709
step: 150, loss: 0.04446474462747574
step: 160, loss: 0.013548307120800018
step: 170, loss: 0.003154299920424819
step: 180, loss: 0.007206766866147518
step: 190, loss: 0.00668864231556654
```

```
step: 200, loss: 0.005325923673808575
step: 210, loss: 0.001408223994076252
step: 220, loss: 0.004516588523983955
step: 230, loss: 0.002003903966397047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932432432432432, f1=0.9831271091113611,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06018170714378357
step: 10, loss: 0.004284980241209269
step: 20, loss: 0.005751398857682943
step: 30, loss: 0.00223508570343256
step: 40, loss: 0.17445801198482513
step: 50, loss: 0.0034544591326266527
step: 60, loss: 0.00156334624625742444
step: 70, loss: 0.0562438927590847
step: 80, loss: 0.0021757916547358036
step: 90, loss: 0.0685972198843956
step: 100, loss: 0.005125819705426693
step: 110, loss: 0.001760714570991695
step: 120, loss: 0.018832100555300713
step: 130, loss: 0.025499563664197922
step: 140, loss: 0.0010994179174304008
step: 150, loss: 0.0011676237918436527
step: 160, loss: 0.0011345319217070937
step: 170, loss: 0.021641427651047707
step: 180, loss: 0.1604745239019394
step: 190, loss: 0.004682186059653759
step: 200, loss: 0.006377057172358036
step: 210, loss: 0.008736466988921165
step: 220, loss: 0.0010647289454936981
step: 230, loss: 0.009996030479669571
/usr/local/lib/python3.7/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=0.9864864864864865, f1=0.9784335981838819,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0008165739127434790
step: 10, loss: 0.0010328072821721435
step: 20, loss: 0.0024474202655255795
step: 30, loss: 0.0016303748125210404
step: 40, loss: 0.04222530871629715
step: 50, loss: 0.0011275263968855143
step: 60, loss: 0.001429474214091897
step: 70, loss: 0.0011112606152892113
step: 80, loss: 0.11359497904777527
step: 90, loss: 0.08874198794364929
step: 100, loss: 0.007081211544573307
step: 110, loss: 0.007878576405346394
step: 120, loss: 0.0013872194103896618
step: 130, loss: 0.0010245456360280514
step: 140, loss: 0.001591815846040845
step: 150, loss: 0.08949631452560425
step: 160, loss: 0.001309212064370513
step: 170, loss: 0.009941913187503815
step: 180, loss: 0.0007437823223881423
step: 190, loss: 0.03203548118472099
step: 200, loss: 0.001350303879007697
step: 210, loss: 0.017385894432663918
step: 220, loss: 0.0008558028494007885
step: 230, loss: 0.0019764872267842293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9865470852017937, f1=0.9864559819413092,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003755402285605669
step: 10, loss: 0.001630302518606186
step: 20, loss: 0.0018657209584489465
step: 30, loss: 0.0009411876671947539
step: 40, loss: 0.00047141656978055835
step: 50, loss: 0.000540331588126719
step: 60, loss: 0.0016033831052482128
step: 70, loss: 0.0014360372442752123
step: 80, loss: 0.02995665743947029
step: 90, loss: 0.009731543250381947
step: 100, loss: 0.0006116524455137551
step: 110, loss: 0.03777105733752251
step: 120, loss: 0.0003266755666118115
step: 130, loss: 0.0008280631736852229
step: 140, loss: 0.00040156865725293756
step: 150, loss: 0.00011921800614800304
```

```
step: 160, loss: 0.001121966284699738
step: 170, loss: 0.0002790651051327586
step: 180, loss: 0.0005813130992464721
step: 190, loss: 0.0002984962775371969
step: 200, loss: 0.003246329491958022
step: 210, loss: 0.0009861374273896217
step: 220, loss: 0.004693740047514439
step: 230, loss: 0.0014066604198887944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9887892376681614, f1=0.9821428571428571,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003553216578438878
step: 10, loss: 0.0011366379912942648
step: 20, loss: 0.0006431720103137195
step: 30, loss: 0.00041219525155611336
step: 40, loss: 0.0035035477485507727
step: 50, loss: 0.0010910218115895987
step: 60, loss: 0.0005091644125059247
step: 70, loss: 0.0005002508987672627
step: 80, loss: 0.00046429995563812554
step: 90, loss: 0.006758709438145161
step: 100, loss: 0.0004532508901320398
step: 110, loss: 0.0006726842257194221
step: 120, loss: 0.0019079305930063128
step: 130, loss: 0.015104400925338268
step: 140, loss: 0.0015531278913840652
step: 150, loss: 0.010790293104946613
step: 160, loss: 0.0024936210829764605
step: 170, loss: 0.0002900583494949202836
step: 180, loss: 0.0007747969939373434
step: 190, loss: 0.0003261083329562098
step: 200, loss: 0.00033139323932118714
step: 210, loss: 0.0004373075207695365
step: 220, loss: 0.0007859364850446582
step: 230, loss: 0.0002860899840015918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9888392857142857, f1=0.9854423292273236,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0004505910037551075
step: 10, loss: 0.0012464930769056082
step: 20, loss: 0.0007618993986397982
step: 30, loss: 0.0003297958173789084
step: 40, loss: 0.00037391376099549234
step: 50, loss: 0.0007571210153400898
step: 60, loss: 0.00041587225859984756
step: 70, loss: 0.0001156710204668343
step: 80, loss: 0.06050609052181244
step: 90, loss: 0.0005036108195781708
step: 100, loss: 0.00033445903682149947
step: 110, loss: 0.010804558172821999
step: 120, loss: 0.010107372887432575
step: 130, loss: 0.013568954542279243
step: 140, loss: 0.0010722920997068286
step: 150, loss: 0.1068301573395729
step: 160, loss: 0.005105364602059126
step: 170, loss: 0.008553990162909031
step: 180, loss: 0.002621295163407922
step: 190, loss: 0.0014044198906049132
step: 200, loss: 0.00990312546491623
step: 210, loss: 0.00023654733377043158
step: 220, loss: 0.00013050998677499592
step: 230, loss: 0.00011195051047252491
/usr/local/lib/python3.7/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=0.9887133182844244, f1=0.976054732041049,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.154453876661137e-05
step: 10, loss: 0.00012148282257840037
step: 20, loss: 0.00013000860053580254
step: 30, loss: 0.04939940944314003
step: 40, loss: 0.00013172214676160365
step: 50, loss: 0.0012791329063475132
step: 60, loss: 0.09099627286195755
step: 70, loss: 0.035944946110248566
step: 80, loss: 0.0005687402444891632
step: 90, loss: 0.0715271383523941
step: 100, loss: 7.353016553679481e-05
step: 110, loss: 0.0001032376749208197
step: 120, loss: 0.00010523723176447675
step: 130, loss: 0.00043652651947923005
step: 140, loss: 0.0014683542540296912
step: 150, loss: 0.0005273779388517141
```

```
step: 160, loss: 0.0087157646194410038
step: 170, loss: 5.1513394282665104e-05
step: 180, loss: 0.0002869571908377111
step: 190, loss: 4.906229878542945e-05
step: 200, loss: 0.00010156539792660624
step: 210, loss: 0.010731658898293972
step: 220, loss: 0.0007659607217647135
step: 230, loss: 0.00066206627525358919
```
/usr/local/lib/python3.7/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=0.9886877828054299, f1=0.9818594104308391,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```
step: 0, loss: 0.00012463687744457275
step: 10, loss: 0.0001911931176437065
step: 20, loss: 0.00017465389100834727
step: 30, loss: 4.0491293475497514e-05
step: 40, loss: 0.000594909826759249
step: 50, loss: 5.2150488045299426e-05
step: 60, loss: 5.719036198570393e-05
step: 70, loss: 0.00010736445983638987
step: 80, loss: 3.247877975809388e-05
step: 90, loss: 3.824110171990469e-05
step: 100, loss: 4.325882764533162e-05
step: 110, loss: 4.0227263525594026e-05
step: 120, loss: 3.587211176636629e-05
step: 130, loss: 0.0006704847910441458
step: 140, loss: 3.6538407584885135e-05
step: 150, loss: 7.849592657294124e-05
step: 160, loss: 2.571077857282944e-05
step: 170, loss: 4.531459853751585e-05
step: 180, loss: 0.0011113606160506606
step: 190, loss: 0.00020753240096382797
step: 200, loss: 5.3305058827390894e-05
step: 210, loss: 0.0013697142712771893
step: 220, loss: 0.10806392133235931
step: 230, loss: 4.763343167724088e-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=0.9841628959276018, f1=0.9806598407281,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```
step: 0, loss: 5.3254145313985646e-05
step: 10, loss: 0.000555530481506139
step: 20, loss: 0.00018885220924858004
step: 30, loss: 0.0002283869544044137
step: 40, loss: 2.6924881240120158e-05
step: 50, loss: 8.66721456986852e-05
step: 60, loss: 0.03916575759649277
step: 70, loss: 0.0031546929385513067
step: 80, loss: 0.000633632589597255
step: 90, loss: 0.0004967499407939613
step: 100, loss: 0.00010107297566719353
step: 110, loss: 0.005658675916492939
step: 120, loss: 0.0012223976664245129
step: 130, loss: 4.954711766913533e-05
step: 140, loss: 0.009554238058626652
step: 150, loss: 0.00011458644439699128
step: 160, loss: 5.514559961738996e-05
step: 170, loss: 7.547323184553534e-05
step: 180, loss: 5.452285768114962e-05
step: 190, loss: 5.3429848776431754e-05
step: 200, loss: 0.00039976846892386675
step: 210, loss: 0.00011754554725484923
step: 220, loss: 0.0003377509710844606
step: 230, loss: 0.00020727851369883865
/usr/local/lib/python3.7/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=0.9886877828054299, f1=0.984090909090909,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013789156218990684
step: 10, loss: 7.327013008762151e-05
step: 20, loss: 0.0031370585784316063
step: 30, loss: 0.030279049649834633
step: 40, loss: 0.00046466183266602457
step: 50, loss: 0.0010532819433137774
step: 60, loss: 0.000132107685203664
step: 70, loss: 0.00011960107804043218
step: 80, loss: 2.90442712866934e-05
step: 90, loss: 4.708743290393613e-05
step: 100, loss: 4.316902413847856e-05
step: 110, loss: 5.9131554735358804e-05
step: 120, loss: 7.762986206216738e-05
step: 130, loss: 6.21126891928725e-05
step: 140, loss: 5.3710362408310175e-05
step: 150, loss: 0.0002814010949805379
```

```
step: 160, loss: 0.00017870862211566418
step: 170, loss: 0.0034226577263325453
step: 180, loss: 0.00417589396238327
step: 190, loss: 0.0031752826180309057
step: 200, loss: 0.008803204633295536
step: 210, loss: 0.0014561752323061228
step: 220, loss: 0.007453982252627611
step: 230, loss: 0.0002533521910663694
/usr/local/lib/python3.7/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=0.9886877828054299, f1=0.9840546697038726,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010515110334381461
step: 10, loss: 6.440265133278444e-05
step: 20, loss: 0.0009238714701496065
step: 30, loss: 9.05024353414774e-05
step: 40, loss: 0.0008110115304589272
step: 50, loss: 0.0002957353135570884
step: 60, loss: 5.780219362350181e-05
step: 70, loss: 0.00046675826888531446
step: 80, loss: 0.004108409862965345
step: 90, loss: 5.652449181070551e-05
step: 100, loss: 7.05548663972877e-05
step: 110, loss: 0.0003428288910072297
step: 120, loss: 0.00010640952677931637
step: 130, loss: 3.81004429073073e-05
step: 140, loss: 4.779127993970178e-05
step: 150, loss: 4.85468735860195e-05
step: 160, loss: 0.0025726081803441048
step: 170, loss: 5.7591205404605716e-05
step: 180, loss: 0.011750254780054092
step: 190, loss: 6.685370317427441e-05
step: 200, loss: 2.828686046996154e-05
step: 210, loss: 0.0007546636625193059
step: 220, loss: 3.9359263610094786e-05
step: 230, loss: 0.00011213670950382948
/usr/local/lib/python3.7/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=0.9898305084745763, f1=0.9805714285714285,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0001397736486978829
step: 10, loss: 3.2233168894890696e-05
step: 20, loss: 7.285708124982193e-05
step: 30, loss: 7.686640310566872e-05
step: 40, loss: 3.699370427057147e-05
step: 50, loss: 3.1067265808815137e-05
step: 60, loss: 5.183332905289717e-05
step: 70, loss: 8.587232878198847e-05
step: 80, loss: 3.2255076803267e-05
step: 90, loss: 0.0001973076577996835
step: 100, loss: 4.3898882722714916e-05
step: 110, loss: 9.282893006457016e-05
step: 120, loss: 1.670371784712188e-05
step: 130, loss: 8.11149729997851e-05
step: 140, loss: 3.6822337278863415e-05
step: 150, loss: 4.483161319512874e-05
step: 160, loss: 0.00029917003121227026
step: 170, loss: 5.2680632506962866e-05
step: 180, loss: 3.222220766474493e-05
step: 190, loss: 4.5437693188432604e-05
step: 200, loss: 4.9620528443483636e-05
step: 210, loss: 2.5714602088555694e-05
step: 220, loss: 6.107246736064553e-05
step: 230, loss: 5.629779116134159e-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=0.9887640449438202, f1=0.9853438556933484,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00934627465903759
step: 10, loss: 3.405145980650559e-05
step: 20, loss: 0.00012996752047911286
step: 30, loss: 2.979727833007928e-05
step: 40, loss: 2.9092923796270043e-05
step: 50, loss: 2.727164792304393e-05
step: 60, loss: 0.02620982751250267
step: 70, loss: 0.0019814535044133663
step: 80, loss: 3.757430749828927e-05
step: 90, loss: 2.8359227144392207e-05
step: 100, loss: 2.1926261979388073e-05
step: 110, loss: 4.2883391870418563e-05
step: 120, loss: 0.03246847912669182
step: 130, loss: 4.751499727717601e-05
step: 140, loss: 0.001922879251651466
step: 150, loss: 4.6644534450024366e-05
```

```
step: 160, loss: 0.01592571660876274
step: 170, loss: 2.1318881408660673e-05
step: 180, loss: 0.00019573299505282193
step: 190, loss: 0.0007014813018031418
step: 200, loss: 0.00012794544454663992
step: 210, loss: 0.000692116329446435
step: 220, loss: 3.38049721904099e-05
step: 230, loss: 3.921350435120985e-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=0.9898305084745763, f1=0.984090909090909,
best_f1=0.9831271091113611
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, 206.93it/s]
load_f1 = 0.9932432432432432
real_f1 = 0.9932432432432432
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
4398it [00:23, 190.91it/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 roberta \
  --fp16 \
  --save_model \
  --da entry_swap  \
  --da span_shuffle \
  --summarize

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

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

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

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
```

```
loss_scale              : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6399959325790405
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4817136526107788
step: 20, loss: 0.3673175275325775
step: 30, loss: 0.3566621243953705
step: 40, loss: 0.26343798637390137
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 50, loss: 0.31334301829338074
step: 60, loss: 0.2513919174671173
step: 70, loss: 0.21508099138736725
step: 80, loss: 0.14505931735038757
step: 90, loss: 0.11787020415067673
step: 100, loss: 0.3078280985355377
step: 110, loss: 0.15991993248462677
step: 120, loss: 0.09101608395576477
step: 130, loss: 0.2060588002204895
step: 140, loss: 0.12679016590118408
step: 150, loss: 0.05766721069812775
step: 160, loss: 0.11345245689153671
step: 170, loss: 0.0381021723151207
step: 180, loss: 0.05139021575450897
step: 190, loss: 0.039411723613739014
step: 200, loss: 0.0411225780844684
step: 210, loss: 0.16557028889656067
step: 220, loss: 0.09033844619989395
step: 230, loss: 0.09424065053462982
step: 240, loss: 0.033107295632362366
step: 250, loss: 0.06975454092025757
step: 260, loss: 0.08265283703804016
step: 270, loss: 0.24688072502613068
```

```
step: 280, loss: 0.06900816410779953
step: 290, loss: 0.0604005828499794
step: 300, loss: 0.0717972069978714
step: 310, loss: 0.20610181987285614
step: 320, loss: 0.06531716883182526
step: 330, loss: 0.10754071176052094
step: 340, loss: 0.4123305082321167
step: 350, loss: 0.33433640003204346
step: 360, loss: 0.053571175783872604
step: 370, loss: 0.034107036888599396
step: 380, loss: 0.09549641609191895
step: 390, loss: 0.01768111251294613
step: 400, loss: 0.01846632920205593
step: 410, loss: 0.23397159576416016
step: 420, loss: 0.019186832010746002
step: 430, loss: 0.017472805455327034
step: 440, loss: 0.056267548352479935
step: 450, loss: 0.03214958682656288
step: 460, loss: 0.012458994053304195
step: 470, loss: 0.025060927495360374
step: 480, loss: 0.12554244697093964
step: 490, loss: 0.19728365540504456
step: 500, loss: 0.15091049671173096
step: 510, loss: 0.06759326159954071
step: 520, loss: 0.03152482584118843
step: 530, loss: 0.15386748313903809
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9112865768694844, f1=0.9165120593692021,
best_f1=0.9165120593692021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08787281811237335
step: 10, loss: 0.0635317787528038
step: 20, loss: 0.02606472745537758
step: 30, loss: 0.14497850835323334
step: 40, loss: 0.06168085336685181
step: 50, loss: 0.05226801335811615
step: 60, loss: 0.052089739590883255
step: 70, loss: 0.03892073780298233
step: 80, loss: 0.04956533759832382
step: 90, loss: 0.018316011875867844
```

```
step: 100, loss: 0.12417271733283997
step: 110, loss: 0.01644684188067913
step: 120, loss: 0.060294609516859055
step: 130, loss: 0.006843869108706713
step: 140, loss: 0.13027885556221008
step: 150, loss: 0.013176475651562214
step: 160, loss: 0.0452507920563221
step: 170, loss: 0.037929896265268326
step: 180, loss: 0.05911636725068092
step: 190, loss: 0.07053299993276596
step: 200, loss: 0.21449458599090576
step: 210, loss: 0.10094975680112839
step: 220, loss: 0.0011400793446227908
step: 230, loss: 0.1177673190832138
step: 240, loss: 0.04648010805249214
step: 250, loss: 0.07191763073205948
step: 260, loss: 0.05921536311507225
step: 270, loss: 0.0229471605271101
step: 280, loss: 0.051545605063438416
step: 290, loss: 0.033594079315662384
step: 300, loss: 0.03771735355257988
step: 310, loss: 0.016922304406762123
step: 320, loss: 0.01706031896173954
step: 330, loss: 0.05466814711689949
step: 340, loss: 0.057382382452487946
step: 350, loss: 0.0028471527621150017
step: 360, loss: 0.14521074295043945
step: 370, loss: 0.004858094267547131
step: 380, loss: 0.11067219823598862
step: 390, loss: 0.008361770771443844
step: 400, loss: 0.03956260532140732
step: 410, loss: 0.033788442611694336
step: 420, loss: 0.03177362680435181
step: 430, loss: 0.12383686006069183
step: 440, loss: 0.09595087915658951
step: 450, loss: 0.030845537781715393
step: 460, loss: 0.04312979429960251
step: 470, loss: 0.02250378578901291
step: 480, loss: 0.010027151554822922
step: 490, loss: 0.032719384878873825
step: 500, loss: 0.013181556016206741
step: 510, loss: 0.016042808070778847
step: 520, loss: 0.2833297848701477
step: 530, loss: 0.09282325208187103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
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=0.9316596931659693, f1=0.9409576940957695,
best_f1=0.9409576940957695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16037794947624207
step: 10, loss: 0.05516272410750389
step: 20, loss: 0.0050053279455751181
step: 30, loss: 0.09849780052900314
step: 40, loss: 0.012134041637182236
step: 50, loss: 0.02327205054461956
step: 60, loss: 0.0021298248320817947
step: 70, loss: 0.008480089716613293
step: 80, loss: 0.01902315393090248
step: 90, loss: 0.0010781417367979884
step: 100, loss: 0.1556846648454666
step: 110, loss: 0.011283843778073788
step: 120, loss: 0.20610088109970093
step: 130, loss: 0.10269393771886826
step: 140, loss: 0.019205976277589798
step: 150, loss: 0.016495879739522934
step: 160, loss: 0.0246881153434515
step: 170, loss: 0.00884312018752098
step: 180, loss: 0.029784904792904854
step: 190, loss: 0.00844857469201088
step: 200, loss: 0.020617470145225525
step: 210, loss: 0.023138079792261124
step: 220, loss: 0.12175821512937546
step: 230, loss: 0.01880684308707714
step: 240, loss: 0.06256330758333206
step: 250, loss: 0.06375366449356079
step: 260, loss: 0.10624314844608307
step: 270, loss: 0.02106393128633499
step: 280, loss: 0.00338856247253716
step: 290, loss: 0.007134269457310438
step: 300, loss: 0.2020885944366455
step: 310, loss: 0.03615507856011391
step: 320, loss: 0.04663224518299103
step: 330, loss: 0.03338287025690079
step: 340, loss: 0.007074097637087107
step: 350, loss: 0.0685228705406189
step: 360, loss: 0.025189172476530075
step: 370, loss: 0.031112713739275932
step: 380, loss: 0.01460188627243042
step: 390, loss: 0.01374234538525343
```

```
step: 400, loss: 0.15290814638137817
step: 410, loss: 0.09983616322278976
step: 420, loss: 0.010211063548922539
step: 430, loss: 0.06125320866703987
step: 440, loss: 0.1379980742931366
step: 450, loss: 0.06259215623140335
step: 460, loss: 0.1345571130514145
step: 470, loss: 0.023825673386454582
step: 480, loss: 0.08251635730266571
step: 490, loss: 0.10040929168462753
step: 500, loss: 0.025892222300171852
step: 510, loss: 0.04769255593419075
step: 520, loss: 0.00493106571957469
step: 530, loss: 0.003771140007302165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9481065918653577, f1=0.9466791393826004,
best_f1=0.9466791393826004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014988696202635765
step: 10, loss: 0.004952911753207445
step: 20, loss: 0.03300678730010986
step: 30, loss: 0.17656679451465607
step: 40, loss: 0.029027465730905533
step: 50, loss: 0.046416353434324265
step: 60, loss: 0.010388594120740089
step: 70, loss: 0.06094607338309288
step: 80, loss: 0.1271354854106903
step: 90, loss: 0.1345202922821045
step: 100, loss: 0.0036221498157829046
step: 110, loss: 0.05588801950216293
step: 120, loss: 0.014150410890579224
step: 130, loss: 0.006992524955421686
step: 140, loss: 0.0073534478433430195
step: 150, loss: 0.022156132385134697
step: 160, loss: 0.05361786112189293
step: 170, loss: 0.0035692472010850906
step: 180, loss: 0.06555303931236267
step: 190, loss: 0.09306202828884125
step: 200, loss: 0.008444014005362988
step: 210, loss: 0.001113003003410995
```

```
step: 220, loss: 0.014573388732969761
step: 230, loss: 0.01183465588837862
step: 240, loss: 0.011260242201387882
step: 250, loss: 0.14700300991535187
step: 260, loss: 0.0033640023320913315
step: 270, loss: 0.1401236355304718
step: 280, loss: 0.008741648867726326
step: 290, loss: 0.07588399946689606
step: 300, loss: 0.01866648718714714
step: 310, loss: 0.009412775747478008
step: 320, loss: 0.06608735024929047
step: 330, loss: 0.019200438633561134
step: 340, loss: 0.0011296682059764862
step: 350, loss: 0.2216813564300537
step: 360, loss: 0.013261736370623112
step: 370, loss: 0.0071573457680642605
step: 380, loss: 0.0024922711309045553
step: 390, loss: 0.00032237343839369714
step: 400, loss: 0.01441811304539442
step: 410, loss: 0.0035318387672305107
step: 420, loss: 0.006501391064375639
step: 430, loss: 0.003211434232071042
step: 440, loss: 0.02749587781727314
step: 450, loss: 0.02504838816821575
step: 460, loss: 0.0806700736284256
step: 470, loss: 0.004754167515784502
step: 480, loss: 0.007561547681689262
step: 490, loss: 0.00041188651812262833
step: 500, loss: 0.09979434311389923
step: 510, loss: 0.13254964351654053
step: 520, loss: 0.011604289524257183
step: 530, loss: 0.13664349913597107
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9499536607970344, f1=0.945707656612529,
best_f1=0.945707656612529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002511895028874278
step: 10, loss: 0.005307809449732304
step: 20, loss: 0.031125495210289955
step: 30, loss: 0.02505660057067871
```

```
step: 40, loss: 0.0004647540918085724
step: 50, loss: 0.11505823582410812
step: 60, loss: 0.0025011992547661066
step: 70, loss: 0.0005865760613232851
step: 80, loss: 0.0021829756442457438
step: 90, loss: 0.10566878318786621
step: 100, loss: 0.005298294126987457
step: 110, loss: 0.008008946664631367
step: 120, loss: 0.17335213720798492
step: 130, loss: 0.01625499315559864
step: 140, loss: 0.01169678382575512
step: 150, loss: 0.04852130636572838
step: 160, loss: 0.0027525005862116814
step: 170, loss: 0.08982688188552856
step: 180, loss: 0.015437407419085503
step: 190, loss: 0.023884935304522514
step: 200, loss: 0.004195868037641048
step: 210, loss: 0.0032875207252800465
step: 220, loss: 0.0028671808540821075
step: 230, loss: 0.001220781821757555
step: 240, loss: 0.018511733040213585
step: 250, loss: 0.0797935426235199
step: 260, loss: 0.0007197157829068601
step: 270, loss: 0.003417906817048788
step: 280, loss: 0.010108157992362976
step: 290, loss: 0.045491382479667664
step: 300, loss: 0.16999554634094238
step: 310, loss: 0.06174968183040619
step: 320, loss: 0.05256974697113037
step: 330, loss: 0.006500737275928259
step: 340, loss: 0.001698954845778644
step: 350, loss: 0.0003795779775828123
step: 360, loss: 8.473161869915202e-05
step: 370, loss: 0.0010774467373266816
step: 380, loss: 0.0018386380979791284
step: 390, loss: 0.016980644315481186
step: 400, loss: 0.008311874233186245
step: 410, loss: 0.049012377858161926
step: 420, loss: 0.1480141133069992
step: 430, loss: 0.10021655261516571
step: 440, loss: 0.00688263401389122
step: 450, loss: 0.026012377813458443
step: 460, loss: 0.09107908606529236
step: 470, loss: 0.04058399796485901
step: 480, loss: 0.01218192558735609
step: 490, loss: 0.003149665892124176
step: 500, loss: 0.0046118986792862415
step: 510, loss: 0.0010390087263658643
```

```
step: 520, loss: 0.2115931361913681
step: 530, loss: 0.008657210506498814
/usr/local/lib/python3.7/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=0.9441860465116279, f1=0.9403054141601109,
best_f1=0.945707656612529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006902698427438736
step: 10, loss: 0.00104858016129583122
step: 20, loss: 0.005389433819800615
step: 30, loss: 0.0005397861823439598
step: 40, loss: 0.0030681719072163105
step: 50, loss: 0.000801389105618
step: 60, loss: 0.005556963384151459
step: 70, loss: 0.0009122865740209818
step: 80, loss: 0.0034685025457292795
step: 90, loss: 0.0054898951202631
step: 100, loss: 0.14156794548034668
step: 110, loss: 0.002387196058407426
step: 120, loss: 0.0209346991032362
step: 130, loss: 0.0017074765637516975
step: 140, loss: 0.0038187194149941206
step: 150, loss: 0.0003232788876630366
step: 160, loss: 0.20858432352542877
step: 170, loss: 0.0004590929893311113
step: 180, loss: 0.006835987791419029
step: 190, loss: 0.010393867269158363
step: 200, loss: 0.0008605755283497274
step: 210, loss: 0.0022612190805375576
step: 220, loss: 0.009042925201356411
step: 230, loss: 0.0028272671625018121
step: 240, loss: 0.12970741093158722
step: 250, loss: 0.01091458648443222
step: 260, loss: 0.0011519192485138774
step: 270, loss: 0.003569537540897727
step: 280, loss: 0.09235893934965134
step: 290, loss: 0.0042412267066653833
step: 300, loss: 0.017238371074199677
step: 310, loss: 0.032495107501745224
step: 320, loss: 0.000131946086185053
step: 330, loss: 0.0007994179031811655
step: 340, loss: 0.0025670661125332117
step: 350, loss: 0.001796691445633769
step: 360, loss: 0.05068305507302284
step: 370, loss: 0.0019293243531137705
```

```
step: 380, loss: 0.00022612478642258793
step: 390, loss: 0.003984046634286642
step: 400, loss: 0.005313800647854805
step: 410, loss: 0.019463760778307915
step: 420, loss: 0.019490454345941544
step: 430, loss: 0.00027466402389109135
step: 440, loss: 0.005059997085481882
step: 450, loss: 0.15999646484851837
step: 460, loss: 0.0004998238873668015
step: 470, loss: 0.0006887381314300001
step: 480, loss: 0.021660298109054565
step: 490, loss: 0.010371607728302479
step: 500, loss: 0.02319067157804966
step: 510, loss: 0.24221089482307434
step: 520, loss: 0.0008244201308116317
step: 530, loss: 0.013598519377410412
/usr/local/lib/python3.7/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=0.9435370975268316, f1=0.9466292134831461,
best_f1=0.945707656612529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01322998572140932
step: 10, loss: 0.0013124952092766762
step: 20, loss: 0.0023338503669947386
step: 30, loss: 0.002951962873339653
step: 40, loss: 0.005773280281573534
step: 50, loss: 0.013630497269332409
step: 60, loss: 0.003849504981189966
step: 70, loss: 0.0019117828924208880
step: 80, loss: 0.0003417681436985731
step: 90, loss: 0.0012431873474270105
step: 100, loss: 0.07878867536783218
step: 110, loss: 0.0007503090891987085
step: 120, loss: 0.0008109008194878697
step: 130, loss: 0.00012646733375731856
step: 140, loss: 0.001918959547765553
step: 150, loss: 0.00026290080859325826
step: 160, loss: 0.00014890111924614757
step: 170, loss: 0.003137258579954505
step: 180, loss: 0.0013457121094688773
step: 190, loss: 0.012722581624984741
step: 200, loss: 0.00047478373744525015
step: 210, loss: 0.00019532606529537588
step: 220, loss: 0.000589648843742907
step: 230, loss: 0.000801604415755719
```

```
step: 240, loss: 0.004405127372592688
step: 250, loss: 0.004229493904858828
step: 260, loss: 0.007076353766024113
step: 270, loss: 0.0011416884372010827
step: 280, loss: 0.005086677148938179
step: 290, loss: 0.02690061926841736
step: 300, loss: 0.0009390267077833414
step: 310, loss: 0.0005760795902460814
step: 320, loss: 0.0009997003944590688
step: 330, loss: 0.00016376159328501672
step: 340, loss: 0.00035796474548988044
step: 350, loss: 0.0003492967807687819
step: 360, loss: 0.0065793609246611595
step: 370, loss: 0.09535159170627594
step: 380, loss: 0.004527245182543993
step: 390, loss: 0.007688415236771107
step: 400, loss: 0.04605437442660332
step: 410, loss: 0.007824222557246685
step: 420, loss: 0.011998475529253483
step: 430, loss: 0.0057731266133487225
step: 440, loss: 0.00034038262674584985
step: 450, loss: 0.08854266256093979
step: 460, loss: 0.0357261598110199
step: 470, loss: 0.058711938560009
step: 480, loss: 0.01187966950237751
step: 490, loss: 0.006253503262996674
step: 500, loss: 0.004345512483268976
step: 510, loss: 0.003144020214676857
step: 520, loss: 0.0007207469898276031
step: 530, loss: 0.005651258397847414
/usr/local/lib/python3.7/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=0.9455719557195572, f1=0.9418282548476454,
best_f1=0.945707656612529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030879899859428406
step: 10, loss: 0.01450889091938734
step: 20, loss: 0.04645409435033798
step: 30, loss: 0.0027505725156515837
step: 40, loss: 0.00019438008894212544
step: 50, loss: 0.00048429641174152493
step: 60, loss: 0.007214342709630728
step: 70, loss: 0.04406515881419182
step: 80, loss: 0.0028949000407010317
step: 90, loss: 0.0012867235345765948
```

```
step: 100, loss: 0.0006993332644924521
step: 110, loss: 0.0002960275742225349
step: 120, loss: 0.003915624227374792
step: 130, loss: 0.0008147390908561647
step: 140, loss: 0.0005628593498840928
step: 150, loss: 0.0016333060339093208
step: 160, loss: 0.00026942635304294527
step: 170, loss: 0.002291975773573359
step: 180, loss: 0.0003453380486462265
step: 190, loss: 0.008551348932087421
step: 200, loss: 0.001878888695500791
step: 210, loss: 0.02370135858654976
step: 220, loss: 0.00010276750253979117
step: 230, loss: 0.008672712370753288
step: 240, loss: 0.011035538278520107
step: 250, loss: 4.1153580241370946e-05
step: 260, loss: 0.00010300760914105922
step: 270, loss: 0.010511746630072594
step: 280, loss: 0.02526702918112278
step: 290, loss: 0.00033372806501574814
step: 300, loss: 5.6101060181390494e-05
step: 310, loss: 0.002264182548969984
step: 320, loss: 0.00021149330132175237
step: 330, loss: 0.0002667135850060731
step: 340, loss: 0.01254034973680973
step: 350, loss: 0.00533700454980135
step: 360, loss: 0.006359475199133158
step: 370, loss: 0.02053310535481407
step: 380, loss: 0.00156157196033746
step: 390, loss: 0.0027975854463875294
step: 400, loss: 0.0033612640108913183
step: 410, loss: 0.005313090048730373
step: 420, loss: 0.0002708693500608206
step: 430, loss: 0.012729438953101635
step: 440, loss: 0.0036718035116791725
step: 450, loss: 0.0006386699387803674
step: 460, loss: 0.0012959132436662912
step: 470, loss: 0.057983674108982086
step: 480, loss: 0.0010368338553234935
step: 490, loss: 0.14652980864048004
step: 500, loss: 0.0006145682418718934
step: 510, loss: 0.0024571334943175316
step: 520, loss: 0.0001874744484666735
step: 530, loss: 0.0004823886847589165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
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=0.9511854951185496, f1=0.949041608228144,
best_f1=0.949041608228144
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026500727981328964
step: 10, loss: 0.002758059184998274
step: 20, loss: 0.0004960749065503478
step: 30, loss: 0.13157010078430176
step: 40, loss: 0.038458529859781265
step: 50, loss: 0.00045227116788737476
step: 60, loss: 0.002650774782523513
step: 70, loss: 0.005447457078844309
step: 80, loss: 0.0023659111466258764
step: 90, loss: 0.04441533610224724
step: 100, loss: 0.00023528291785623878
step: 110, loss: 0.001769129536114633
step: 120, loss: 0.0005344605888240039
step: 130, loss: 7.987237040651962e-05
step: 140, loss: 7.264950545504689e-05
step: 150, loss: 0.0795372799038887
step: 160, loss: 0.00015962497855070978
step: 170, loss: 0.00034366146428510547
step: 180, loss: 0.005190650932490826
step: 190, loss: 0.0029192729853093624
step: 200, loss: 0.00012141622573835775
step: 210, loss: 0.0012222788063809276
step: 220, loss: 0.0004107674758415669
step: 230, loss: 0.0003107994270976633
step: 240, loss: 0.0004894494195468724
step: 250, loss: 0.00014188872592058033
step: 260, loss: 0.0002383508690400049
step: 270, loss: 0.0002551477518863976
step: 280, loss: 0.022066740319132805
step: 290, loss: 0.00012207204417791218
step: 300, loss: 0.00023129665351007134
step: 310, loss: 0.0024333202745765448
step: 320, loss: 0.0031946063973009586
step: 330, loss: 7.339326111832634e-05
step: 340, loss: 0.0010606361320242286
step: 350, loss: 0.02307564951479435
step: 360, loss: 0.02633092552423477
step: 370, loss: 0.00017491122270772457
step: 380, loss: 0.00019284646259620786
step: 390, loss: 2.468649654474575e-05
```

```
step: 400, loss: 0.10434519499540329
step: 410, loss: 0.00040498157613910735
step: 420, loss: 0.000993831199593842
step: 430, loss: 0.05949584022164345
step: 440, loss: 0.00025870994431898
step: 450, loss: 0.0036265039816498756
step: 460, loss: 0.00020604536985047162
step: 470, loss: 9.651430445956066e-05
step: 480, loss: 0.00013308797497302294
step: 490, loss: 0.015141069889068604
step: 500, loss: 0.00020375123131088912
step: 510, loss: 0.04570520296692848
step: 520, loss: 0.015743065625429153
step: 530, loss: 0.0018827663734555244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9537166900420756, f1=0.9541627689429373,
best_f1=0.9541627689429373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007660880801267922
step: 10, loss: 8.191678352886811e-05
step: 20, loss: 7.932903827168047e-05
step: 30, loss: 0.015489267185330391
step: 40, loss: 7.622387784067541e-05
step: 50, loss: 0.000190842998563312
step: 60, loss: 0.0035043784882873297
step: 70, loss: 3.534380084602162e-05
step: 80, loss: 0.0073789688758552074
step: 90, loss: 0.00809003971517086
step: 100, loss: 0.0001058891139109619
step: 110, loss: 0.01108636986464262
step: 120, loss: 0.00012236055044922978
step: 130, loss: 5.812278322991915e-05
step: 140, loss: 5.253671406535432e-05
step: 150, loss: 0.00011283308413112536
step: 160, loss: 0.030392328277230263
step: 170, loss: 0.00015283025277312845
step: 180, loss: 0.00147542217746377
step: 190, loss: 4.547499702312052e-05
step: 200, loss: 0.00010419407772133127
step: 210, loss: 0.014024426229298115
```

```
step: 220, loss: 0.00014314775762613863
step: 230, loss: 4.483537850319408e-05
step: 240, loss: 4.984898987459019e-05
step: 250, loss: 0.0069441720843315125
step: 260, loss: 0.008443271741271019
step: 270, loss: 0.0004948207642883062
step: 280, loss: 0.00033516393159516156
step: 290, loss: 0.00028421636670827866
step: 300, loss: 0.006908648647367954
step: 310, loss: 0.03467368707060814
step: 320, loss: 0.0028298040851950645
step: 330, loss: 4.9673206376610324e-05
step: 340, loss: 0.00021294725593179464
step: 350, loss: 0.040320299565792084
step: 360, loss: 0.0001257346011698246
step: 370, loss: 0.00128216075245291
step: 380, loss: 0.0005216972786001861
step: 390, loss: 0.0015245437389239669
step: 400, loss: 0.021156348288059235
step: 410, loss: 0.005277540069073439
step: 420, loss: 0.00013613792543765157
step: 430, loss: 0.00018969812663272023
step: 440, loss: 0.00018403843569103628
step: 450, loss: 0.0009457519627176225
step: 460, loss: 0.00012662498920690268
step: 470, loss: 0.004727686755359173
step: 480, loss: 0.0010586186544969678
step: 490, loss: 0.07869065552949905
step: 500, loss: 0.00361781963147223
step: 510, loss: 7.436508167302236e-05
step: 520, loss: 0.001724533154629171
step: 530, loss: 0.002551176119595766
/usr/local/lib/python3.7/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=0.9490084985835694, f1=0.9479117785077428,
best_f1=0.9541627689429373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001124336151406169
step: 10, loss: 9.646898979553953e-05
step: 20, loss: 0.00347344484180212
step: 30, loss: 5.41943809366785e-05
step: 40, loss: 0.00013034549192525446
step: 50, loss: 6.722156831528991e-05
step: 60, loss: 5.362757292459719e-05
step: 70, loss: 0.014046486467123032
```

```
step: 80, loss: 5.347406113287434e-05
step: 90, loss: 8.504725701641291e-05
step: 100, loss: 8.5663610661868e-05
step: 110, loss: 0.0003501664032228291
step: 120, loss: 0.0001565219572512433
step: 130, loss: 4.8961628635879606e-05
step: 140, loss: 2.5431743779336102e-05
step: 150, loss: 0.19418524205684662
step: 160, loss: 0.00018952382379211485
step: 170, loss: 6.161383498692885e-05
step: 180, loss: 5.029895328334533e-05
step: 190, loss: 0.00011244387133046985
step: 200, loss: 0.00018031088984571397
step: 210, loss: 3.9802809624234214e-05
step: 220, loss: 0.005719391163438559
step: 230, loss: 2.6105348297278397e-05
step: 240, loss: 0.004429501481354237
step: 250, loss: 0.00019609363516792655
step: 260, loss: 0.00023010160657577217
step: 270, loss: 0.0012721098028123379
step: 280, loss: 0.0062152850441634655
step: 290, loss: 0.0002028077724389732
step: 300, loss: 0.00011627995991148055
step: 310, loss: 0.01010893750935793
step: 320, loss: 0.0001279195857932791
step: 330, loss: 5.875957140233368e-05
step: 340, loss: 0.001736655831336975
step: 350, loss: 4.134108894504607e-05
step: 360, loss: 0.00015486417396459728
step: 370, loss: 8.21464418550022e-05
step: 380, loss: 7.375870336545631e-05
step: 390, loss: 2.223887713626027e-05
step: 400, loss: 0.018677804619073868
step: 410, loss: 0.00047497168998233974
step: 420, loss: 3.820510028162971e-05
step: 430, loss: 0.00035899836802855134
step: 440, loss: 0.00014427112182602286
step: 450, loss: 0.0009026534971781075
step: 460, loss: 0.00018813495989888906
step: 470, loss: 0.0003644039388746023
step: 480, loss: 0.00012319005327299237
step: 490, loss: 3.773875505430624e-05
step: 500, loss: 0.0007787301437929273
step: 510, loss: 0.09663034230470657
step: 520, loss: 7.38126618671231e-05
step: 530, loss: 0.0044329515658319
/usr/local/lib/python3.7/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=0.9506346967559944, f1=0.9441052137153593,
best_f1=0.9541627689429373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.08753364556469e-05
step: 10, loss: 0.026178158819675446
step: 20, loss: 0.009404287673532963
step: 30, loss: 3.2626066968077794e-05
step: 40, loss: 0.0011996752582490444
step: 50, loss: 4.315685873734765e-05
step: 60, loss: 0.00010127252608072013
step: 70, loss: 0.000527353142388165
step: 80, loss: 0.09393537044525146
step: 90, loss: 0.00045246712397783995
step: 100, loss: 0.0007991984602995217
step: 110, loss: 7.204303256003186e-05
step: 120, loss: 0.00017797178588807583
step: 130, loss: 0.0006346068694256246
step: 140, loss: 0.00022425802308134735
step: 150, loss: 0.0005619448493234813
step: 160, loss: 7.046604878269136e-05
step: 170, loss: 5.939240145380609e-05
step: 180, loss: 0.00017864046094473451
step: 190, loss: 0.00010845748329302296
step: 200, loss: 9.641848737373948e-05
step: 210, loss: 7.916627509985119e-05
step: 220, loss: 2.314770063094329e-05
step: 230, loss: 0.0002870018361136317
step: 240, loss: 0.0005196359707042575
step: 250, loss: 0.00570681830868125
step: 260, loss: 2.5166789782815613e-05
step: 270, loss: 0.007731721270829439
step: 280, loss: 0.0001270169741474092
step: 290, loss: 0.00010144760744879022
step: 300, loss: 5.8104815252590925e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 310, loss: 3.62097671313677e-05
step: 320, loss: 0.0002838285290636122
step: 330, loss: 0.0022926167584955692
step: 340, loss: 0.0011482934933155775
step: 350, loss: 5.725803202949464e-05
step: 360, loss: 9.292181493947282e-05
step: 370, loss: 0.0013535460457205772
step: 380, loss: 0.00021012048819102347
step: 390, loss: 2.2295220333035104e-05
step: 400, loss: 1.2185187188151758e-05
```

```
step: 410, loss: 0.0003100537578575313
step: 420, loss: 6.407393811969087e-05
step: 430, loss: 0.00188731262460351
step: 440, loss: 0.0001951288286363706
step: 450, loss: 4.427629755809903e-05
step: 460, loss: 0.0001478145713917911
step: 470, loss: 0.0004669697373174131
step: 480, loss: 0.0042157103307545185
step: 490, loss: 4.0615930629428476e-05
step: 500, loss: 1.5574260032735765e-05
step: 510, loss: 7.07491853972897e-05
step: 520, loss: 4.837513915845193e-05
step: 530, loss: 1.506106218585046e-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=0.9503480278422275, f1=0.9489795918367347,
best_f1=0.9541627689429373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.223266038636211e-06
step: 10, loss: 3.619578274083324e-05
step: 20, loss: 5.263061757432297e-05
step: 30, loss: 6.619781288463855e-06
step: 40, loss: 0.0017499277601018548
step: 50, loss: 8.052286284510046e-05
step: 60, loss: 0.00022116674517747015
step: 70, loss: 1.2095765669073444e-05
step: 80, loss: 0.006441181991249323
step: 90, loss: 5.790368595626205e-05
step: 100, loss: 4.4615688238991424e-05
step: 110, loss: 5.5133950809249654e-06
step: 120, loss: 0.00014286689111031592
step: 130, loss: 2.2851811081636697e-05
step: 140, loss: 1.993715159187559e-05
step: 150, loss: 2.061824852717109e-05
step: 160, loss: 2.497314926586114e-05
step: 170, loss: 0.00014236436982173473
step: 180, loss: 2.056270204775501e-05
step: 190, loss: 3.0044599043321796e-05
step: 200, loss: 2.0286614017095417e-05
step: 210, loss: 2.2420246750698425e-05
step: 220, loss: 1.2233507732162252e-05
step: 230, loss: 8.06252719485201e-05
step: 240, loss: 2.179915099986829e-05
step: 250, loss: 0.00014869784354232252
step: 260, loss: 1.5943438484100625e-05
```

```
step: 270, loss: 1.4151813957141712e-05
step: 280, loss: 3.8716327253496274e-05
step: 290, loss: 6.869355729577364e-06
step: 300, loss: 1.0866474440263119e-05
step: 310, loss: 1.8737047867034562e-05
step: 320, loss: 1.4855700101179536e-05
step: 330, loss: 5.0399765314068645e-05
step: 340, loss: 3.8804977521067485e-05
step: 350, loss: 0.00020889524603262544
step: 360, loss: 0.0002972473739646375
step: 370, loss: 8.094957593129948e-06
step: 380, loss: 0.0034415568225085735
step: 390, loss: 0.0005482006235979497
step: 400, loss: 4.848559547099285e-05
step: 410, loss: 1.2293223335291259e-05
step: 420, loss: 9.793629942578264e-06
step: 430, loss: 0.00010867384116863832
step: 440, loss: 4.762732351082377e-05
step: 450, loss: 1.9154715118929744e-05
step: 460, loss: 0.000306666741380468
step: 470, loss: 8.504503966832999e-06
step: 480, loss: 4.298948624636978e-06
step: 490, loss: 2.4088249119813554e-05
step: 500, loss: 2.8881044272566214e-05
step: 510, loss: 0.00010787845531012863
step: 520, loss: 0.0010603530099615455
step: 530, loss: 0.00019864164642058313
/usr/local/lib/python3.7/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=0.9500462534690102, f1=0.9468822170900694,
best_f1=0.9541627689429373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.395155211677775e-05
step: 10, loss: 2.1776968424092047e-05
step: 20, loss: 3.842139267362654e-05
step: 30, loss: 0.0003373397048562765
step: 40, loss: 1.097443873732118e-05
step: 50, loss: 6.41666556475684e-05
step: 60, loss: 0.00019268292817287147
step: 70, loss: 1.2013765626761597e-05
step: 80, loss: 4.536937194643542e-05
step: 90, loss: 7.640463991265278e-06
step: 100, loss: 3.087875302298926e-05
step: 110, loss: 4.225597876938991e-05
step: 120, loss: 3.93109476135578e-05
```

```
step: 130, loss: 8.042829904297832e-06
step: 140, loss: 0.00016386511560995132
step: 150, loss: 0.0004360891762189567
step: 160, loss: 8.263713971246034e-05
step: 170, loss: 0.04323684796690941
step: 180, loss: 3.1577572372043505e-05
step: 190, loss: 0.00010467538959346712
step: 200, loss: 0.00013342789316084236
step: 210, loss: 6.754677451681346e-05
step: 220, loss: 1.314229939453071e-05
step: 230, loss: 0.00017895993369165808
step: 240, loss: 3.697179272421636e-05
step: 250, loss: 0.0037641641683876514
step: 260, loss: 9.99638214125298e-05
step: 270, loss: 0.0004583548870868981
step: 280, loss: 7.307979103643447e-05
step: 290, loss: 7.484596426365897e-05
step: 300, loss: 1.4214931779950052e-05
step: 310, loss: 0.00015330093447118998
step: 320, loss: 1.672867438173853e-05
step: 330, loss: 0.0014731206465512514
step: 340, loss: 2.2122760128695518e-05
step: 350, loss: 4.258969784132205e-05
step: 360, loss: 0.0002550509525462985
step: 370, loss: 3.272304093115963e-05
step: 380, loss: 0.0019462152849882841
step: 390, loss: 1.727222843328491e-05
step: 400, loss: 0.005403292365372181
step: 410, loss: 1.1447556062194053e-05
step: 420, loss: 9.220344509230927e-05
step: 430, loss: 0.000502042646985501
step: 440, loss: 2.780324030027259e-05
step: 450, loss: 1.2699256330961362e-05
step: 460, loss: 0.01006422471255064
step: 470, loss: 7.934797395137139e-06
step: 480, loss: 7.864003237045836e-06
step: 490, loss: 6.068463335395791e-06
step: 500, loss: 2.9328724849619903e-05
step: 510, loss: 0.000707273546140641
step: 520, loss: 6.8171898419677746e-06
step: 530, loss: 1.7864631445263512e-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=0.9523809523809524, f1=0.9456419868791003,
best_f1=0.9541627689429373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.6538542695343494e-05
step: 10, loss: 1.3887101886211894e-05
step: 20, loss: 0.0018933919491246343
step: 30, loss: 0.00024241040227934718
step: 40, loss: 1.408472417097073e-05
step: 50, loss: 9.704130206955597e-05
step: 60, loss: 3.2391031709266827e-05
step: 70, loss: 5.7240318710682914e-05
step: 80, loss: 9.752567166287918e-06
step: 90, loss: 6.191396551002981e-06
step: 100, loss: 6.4409823608002625e-06
step: 110, loss: 5.00005335197784e-05
step: 120, loss: 1.0262994692311622e-05
step: 130, loss: 9.726540156407282e-06
step: 140, loss: 6.152927380753681e-05
step: 150, loss: 1.059426722349599e-05
step: 160, loss: 1.2024564057355747e-05
step: 170, loss: 6.9066009018570185e-06
step: 180, loss: 2.105806015606504e-05
step: 190, loss: 0.00010422812192700803
step: 200, loss: 3.147417737636715e-05
step: 210, loss: 1.682253241597209e-05
step: 220, loss: 1.588685881870333e-05
step: 230, loss: 5.249856985756196e-05
step: 240, loss: 1.1034047020075377e-05
step: 250, loss: 1.6754731404944323e-05
step: 260, loss: 4.354848897492047e-06
step: 270, loss: 1.201337272505043e-05
step: 280, loss: 7.2307270784222055e-06
step: 290, loss: 5.712646816391498e-05
step: 300, loss: 7.059350082272431e-06
step: 310, loss: 0.0002100665442412719
step: 320, loss: 4.597121733240783e-05
step: 330, loss: 8.635111043986399e-06
step: 340, loss: 2.277781823067926e-05
step: 350, loss: 0.006798944901674986
step: 360, loss: 8.277519555122126e-06
step: 370, loss: 2.7001806301996112e-05
step: 380, loss: 2.2802747480454855e-05
step: 390, loss: 2.0371160644572228e-05
step: 400, loss: 0.009983277879655361
step: 410, loss: 2.5443940103286877e-05
step: 420, loss: 3.0503750167554244e-05
step: 430, loss: 4.142507350479718e-06
step: 440, loss: 0.00026364304358139634
step: 450, loss: 0.00012473922106437385
step: 460, loss: 0.0008022578549571335
```

```
step: 470, loss: 4.406999778439058e-06
step: 480, loss: 0.0010715401731431484
step: 490, loss: 7.062564691295847e-05
step: 500, loss: 4.839026223635301e-05
step: 510, loss: 5.181850156077417e-06
step: 520, loss: 6.946739449631423e-05
step: 530, loss: 5.695920663129073e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9532710280373832, f1=0.9457436856875585,
best_f1=0.9541627689429373
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, 253.05it/s]
load_f1 = 0.9530956848030018
real_f1 = 0.9531396438612934
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, 196.96it/s]
```

1.1.4   Amazon-Google - Running the matcher

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

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

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

```
Selected optimization level 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: 0.5137475728988647
step: 10, loss: 0.3465289771556854
step: 20, loss: 0.47972792387008667
step: 30, loss: 0.36519330739974976
step: 40, loss: 0.30490341782569885
step: 50, loss: 0.42212793231010437
step: 60, loss: 0.47976627945899963
step: 70, loss: 0.35419395565986633
step: 80, loss: 0.34803512692451477
step: 90, loss: 0.24331727623939514
step: 100, loss: 0.3067838251590729
step: 110, loss: 0.2384915053844452
step: 120, loss: 0.364633709192276
step: 130, loss: 0.2190997451543808
step: 140, loss: 0.45773404836654663
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.2987203598022461
step: 160, loss: 0.46418875455856323
step: 170, loss: 0.19541925191879272
step: 180, loss: 0.31853967905044556
step: 190, loss: 0.43669092655181885
step: 200, loss: 0.26802948117256165
step: 210, loss: 0.33169636130332947

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5232744783306581, f1=0.5612788632326821,
best_f1=0.5612788632326821
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2623995542526245
step: 10, loss: 0.08912684768438339
step: 20, loss: 0.33415964245796204
step: 30, loss: 0.2943341135978699
step: 40, loss: 0.463588684797287
step: 50, loss: 0.19910986721515656
step: 60, loss: 0.24845010042190552
step: 70, loss: 0.19672289490699768
step: 80, loss: 0.3372224271297455
step: 90, loss: 0.16137059032917023
step: 100, loss: 0.43440285325050354
step: 110, loss: 0.2655116617679596
step: 120, loss: 0.13288834691047668
step: 130, loss: 0.1635052114725113
step: 140, loss: 0.10321231931447983
step: 150, loss: 0.22512474656105042
step: 160, loss: 0.08045128732919693
step: 170, loss: 0.2186182588338852
step: 180, loss: 0.19012437760829926
step: 190, loss: 0.2006445974111557
step: 200, loss: 0.08886446803808212
step: 210, loss: 0.1601886749267578
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6, f1=0.5847457627118645, best_f1=0.5847457627118645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10533684492111206
step: 10, loss: 0.0788038820028305
step: 20, loss: 0.26921480894088745
```

```
step: 30, loss: 0.08496731519699097
step: 40, loss: 0.2948317229747772
step: 50, loss: 0.0773104652762413
step: 60, loss: 0.21305334568023682
step: 70, loss: 0.16661430895328522
step: 80, loss: 0.23117166757583618
step: 90, loss: 0.09612179547548294
step: 100, loss: 0.20737333595752716
step: 110, loss: 0.12311820685863495
step: 120, loss: 0.12418323010206223
step: 130, loss: 0.1623794436454773
step: 140, loss: 0.12058605998754501
step: 150, loss: 0.1701750010251999
step: 160, loss: 0.20681490004062653
step: 170, loss: 0.12968087196350098
step: 180, loss: 0.09703008830547333
step: 190, loss: 0.02219654992222786
step: 200, loss: 0.08338181674480438
step: 210, loss: 0.14737246930599213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6206896551724138, f1=0.6265060240963856,
best_f1=0.6265060240963856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.034389261115632057
step: 10, loss: 0.07937045395374298
step: 20, loss: 0.06284651160240173
step: 30, loss: 0.08328470587730408
step: 40, loss: 0.0369599349796772
step: 50, loss: 0.16827619075775146
step: 60, loss: 0.09966544806957245
step: 70, loss: 0.051452286541461945
step: 80, loss: 0.14187544584274292
step: 90, loss: 0.05257264897227287
step: 100, loss: 0.19968953728675842
step: 110, loss: 0.547524094581604
step: 120, loss: 0.15321338176727295
step: 130, loss: 0.190667062997818
step: 140, loss: 0.44870424270629883
step: 150, loss: 0.07136561721563339
step: 160, loss: 0.14850854873657227
```

```
step: 170, loss: 0.08227815479040146
step: 180, loss: 0.06817315518856049
step: 190, loss: 0.10805284231901169
step: 200, loss: 0.10313357412815094
step: 210, loss: 0.25940391421318054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6452830188679246, f1=0.6628131021194604,
best_f1=0.6628131021194604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14015556871891022
step: 10, loss: 0.052424173802137375
step: 20, loss: 0.09751057624816895
step: 30, loss: 0.01671202853322029
step: 40, loss: 0.06977425515651703
step: 50, loss: 0.15996690094470978
step: 60, loss: 0.1801740801181793
step: 70, loss: 0.10348017513751984
step: 80, loss: 0.03325093537569046
step: 90, loss: 0.029134364798665047
step: 100, loss: 0.010226890444755554
step: 110, loss: 0.04266657307744026
step: 120, loss: 0.14586076140403748
step: 130, loss: 0.05694737285375595
step: 140, loss: 0.29542890191078186
step: 150, loss: 0.13906864821910858
step: 160, loss: 0.06986567378044128
step: 170, loss: 0.04136848822236061
step: 180, loss: 0.3054037094116211
step: 190, loss: 0.08302999287843704
step: 200, loss: 0.3183283507823944
step: 210, loss: 0.15350961685180664
/usr/local/lib/python3.7/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=0.6333333333333333, f1=0.5945945945945946,
best_f1=0.6628131021194604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06837836652994156
```

```
step: 10, loss: 0.09102404117584229
step: 20, loss: 0.0935707613825798
step: 30, loss: 0.022045014426112175
step: 40, loss: 0.0647241398692131
step: 50, loss: 0.1105048656463623
step: 60, loss: 0.12507320940494537
step: 70, loss: 0.06496324390172958
step: 80, loss: 0.15685920417308807
step: 90, loss: 0.09895198047161102
step: 100, loss: 0.07539761066436768
step: 110, loss: 0.03875232860445976
step: 120, loss: 0.08808377385139465
step: 130, loss: 0.04529882222414017
step: 140, loss: 0.07507146149873734
step: 150, loss: 0.038139112293720245
step: 160, loss: 0.12006442248821259
step: 170, loss: 0.058401625603437424
step: 180, loss: 0.06941651552915573
step: 190, loss: 0.030269132927060127
step: 200, loss: 0.07340303808450699
step: 210, loss: 0.19730719923973083
/usr/local/lib/python3.7/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=0.6341463414634146, f1=0.6389413988657845,
best_f1=0.6628131021194604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0394180491566658
step: 10, loss: 0.057746198028326035
step: 20, loss: 0.012768886983394623
step: 30, loss: 0.034884028136730194
step: 40, loss: 0.04213246703147888
step: 50, loss: 0.04841542989015579
step: 60, loss: 0.04177229478955269
step: 70, loss: 0.11576706916093826
step: 80, loss: 0.04646489769220352
step: 90, loss: 0.0159321166574955
step: 100, loss: 0.10843997448682785
step: 110, loss: 0.1566828191280365
step: 120, loss: 0.1157589927315712
step: 130, loss: 0.2778974771499634
step: 140, loss: 0.04353617876768112
step: 150, loss: 0.02362126298248768
step: 160, loss: 0.09839214384555817
step: 170, loss: 0.1619095653295517
step: 180, loss: 0.01979415863752365
```

```
step: 190, loss: 0.056862689554691315
step: 200, loss: 0.05751878023147583
step: 210, loss: 0.12432842701673508
/usr/local/lib/python3.7/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=0.6157112526539279, f1=0.631346578366446,
best_f1=0.6628131021194604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07710892707109451
step: 10, loss: 0.015803158283233643
step: 20, loss: 0.0353272370994091
step: 30, loss: 0.18054816126823425
step: 40, loss: 0.06367463618516922
step: 50, loss: 0.003118594642728567
step: 60, loss: 0.01420912891626358
step: 70, loss: 0.034603558480739594
step: 80, loss: 0.02601664513349533
step: 90, loss: 0.04089251905679703
step: 100, loss: 0.016286827623844147
step: 110, loss: 0.007544963154941797
step: 120, loss: 0.15379635989665985
step: 130, loss: 0.01153677050024271
step: 140, loss: 0.054608460515737534
step: 150, loss: 0.06733647733926773
step: 160, loss: 0.08655262738466263
step: 170, loss: 0.1022152379155159
step: 180, loss: 0.07373686134815216
step: 190, loss: 0.07375994324684143
step: 200, loss: 0.057519201189279556
step: 210, loss: 0.10043882578611374
/usr/local/lib/python3.7/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=0.6324110671936759, f1=0.636, best_f1=0.6628131021194604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08396642655134201
step: 10, loss: 0.2561042606830597
step: 20, loss: 0.046171635389328
step: 30, loss: 0.0008484200807288289
step: 40, loss: 0.002991798799484968
step: 50, loss: 0.09155481308698654
step: 60, loss: 0.005337793380022049
step: 70, loss: 0.13318467140197754
```

```
step: 80, loss: 0.09541609138250351
step: 90, loss: 0.037968166172504425
step: 100, loss: 0.028558405116200447
step: 110, loss: 0.0484786331653595
step: 120, loss: 0.11187443882226944
step: 130, loss: 0.11973035335540771
step: 140, loss: 0.10821598768234253
step: 150, loss: 0.0021381538826972246
step: 160, loss: 0.04210362210869789
step: 170, loss: 0.056004952639341354
step: 180, loss: 0.031548794358968735
step: 190, loss: 0.0010610297322273254
step: 200, loss: 0.00962822139263153
step: 210, loss: 0.04564986005425453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6460348162475823, f1=0.6393762183235867,
best_f1=0.6393762183235867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030534613877534866
step: 10, loss: 0.0320853665471077
step: 20, loss: 0.02161344513297081
step: 30, loss: 0.0007380014285445213
step: 40, loss: 0.004522534087300301
step: 50, loss: 0.006604688707739115
step: 60, loss: 0.0017874392215162516
step: 70, loss: 0.012188141234219074
step: 80, loss: 0.024144167080521584
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.021799154579639435
step: 100, loss: 0.04856928437948227
step: 110, loss: 0.0578354187309742
step: 120, loss: 0.07222797721624374
step: 130, loss: 0.014705699868500233
step: 140, loss: 0.013920498080551624
step: 150, loss: 0.05303030088543892
step: 160, loss: 0.005698305554687977
step: 170, loss: 0.004656917881220579
step: 180, loss: 0.021701738238334656
step: 190, loss: 0.19727720320224762
step: 200, loss: 0.0904262438416481
```

step: 210, loss: 0.1933804601430893
/usr/local/lib/python3.7/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=0.6303030303030304, f1=0.6270491803278688,
best_f1=0.6393762183235867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01655540056526661
step: 10, loss: 0.009841544553637505
step: 20, loss: 0.020346658304333687
step: 30, loss: 0.016142357140779495
step: 40, loss: 0.04256089776754379
step: 50, loss: 0.020871803164482117
step: 60, loss: 0.12819097936153412
step: 70, loss: 0.020673265680670738
step: 80, loss: 0.4860425591468811
step: 90, loss: 0.4227660000324249
step: 100, loss: 0.43560242652893066
step: 110, loss: 0.010929787531495094
step: 120, loss: 0.029456133022904396
step: 130, loss: 0.0018451843643561006
step: 140, loss: 0.004710058216005564
step: 150, loss: 0.030300837010145187
step: 160, loss: 0.006830504629760981
step: 170, loss: 0.01446131058037281
step: 180, loss: 0.02373400889337063
step: 190, loss: 0.01205320842564106
step: 200, loss: 0.0035238116979599
step: 210, loss: 0.10387590527534485
/usr/local/lib/python3.7/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=0.6343434343434343, f1=0.628099173553719,
best_f1=0.6393762183235867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002037924015894532
step: 10, loss: 0.008677089586853981
step: 20, loss: 0.017184024676680565
step: 30, loss: 0.06356043368577957
step: 40, loss: 0.03842411935329437
step: 50, loss: 0.00308588077314198
step: 60, loss: 0.0008093977812677622
step: 70, loss: 0.02033042535185814
step: 80, loss: 0.06695052981376648

```
step: 90, loss: 0.00042419726378284395
step: 100, loss: 0.003819200210273266
step: 110, loss: 0.024792583659291267
step: 120, loss: 0.006440544500946999
step: 130, loss: 0.04731450229883194
step: 140, loss: 0.08787665516138077
step: 150, loss: 0.002347056521102786
step: 160, loss: 0.019075846299529076
step: 170, loss: 0.08104152977466583
step: 180, loss: 0.027473317459225655
step: 190, loss: 0.003114394610747695
step: 200, loss: 0.19604580104351044
step: 210, loss: 0.027772074565291405
/usr/local/lib/python3.7/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=0.634508348794063, f1=0.6477272727272727,
best_f1=0.6393762183235867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09061288088560104
step: 10, loss: 0.004754076711833477
step: 20, loss: 0.014757411554455757
step: 30, loss: 0.007529634516686201
step: 40, loss: 0.005764511413872242
step: 50, loss: 0.13639244437217712
step: 60, loss: 0.016628948971629143
step: 70, loss: 0.02604151889681816
step: 80, loss: 0.015608291141688824
step: 90, loss: 0.012117788195610046
step: 100, loss: 0.014815595000982285
step: 110, loss: 0.026591353118419647
step: 120, loss: 0.03289485722780228
step: 130, loss: 0.0017513781785964966
step: 140, loss: 0.0033611077815294266
step: 150, loss: 0.002381153404712677
step: 160, loss: 0.019615760073065758
step: 170, loss: 0.0041436399333179
step: 180, loss: 0.0315268374979496
step: 190, loss: 0.0023218963760882616
step: 200, loss: 0.05335923656821251
step: 210, loss: 0.0005216505378484726
/usr/local/lib/python3.7/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=0.6363636363636365, f1=0.6386233269598471,
best_f1=0.6393762183235867
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027119931764900684
step: 10, loss: 0.012896222062408924
step: 20, loss: 0.22295117378234863
step: 30, loss: 0.007226626854389906
step: 40, loss: 0.0012848784681409597
step: 50, loss: 0.03748415410518646
step: 60, loss: 0.02224843204021454
step: 70, loss: 0.0051452708430588245
step: 80, loss: 0.23796287178993225
step: 90, loss: 0.004518536850810051
step: 100, loss: 0.004883650224655867
step: 110, loss: 0.0530603863298893
step: 120, loss: 0.023685647174715996
step: 130, loss: 0.007848282344639301
step: 140, loss: 0.016474895179271698
step: 150, loss: 0.15411606431007385
step: 160, loss: 0.004478093236684799
step: 170, loss: 0.10066601634025574
step: 180, loss: 0.0005476034712046385
step: 190, loss: 0.01654941961169243
step: 200, loss: 0.007338858675211668
step: 210, loss: 0.003314180998131633
/usr/local/lib/python3.7/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=0.633587786259542, f1=0.6299810246679317,
best_f1=0.6393762183235867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02021704986691475
step: 10, loss: 0.007141047157347202
step: 20, loss: 0.014522290788590908
step: 30, loss: 0.0130464443936944
step: 40, loss: 0.0015005578752607107
step: 50, loss: 0.0032081622630357742
step: 60, loss: 0.020992908626794815
step: 70, loss: 0.00291714887134730
step: 80, loss: 0.005431440658867359
step: 90, loss: 0.0012149413814768195
step: 100, loss: 0.0017518849344924092
step: 110, loss: 0.10249052196741104
step: 120, loss: 0.004445141647011042
step: 130, loss: 0.005575754679739475
step: 140, loss: 0.0010280997958034277
```

```
step: 150, loss: 0.12227730453014374
step: 160, loss: 0.010828453116118908
step: 170, loss: 0.010719832964241505
step: 180, loss: 0.007661606650799513
step: 190, loss: 0.0005556119722314179
step: 200, loss: 0.0007316162809729576
step: 210, loss: 0.01969262771308422
/usr/local/lib/python3.7/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=0.640625, f1=0.6330097087378641, best_f1=0.6393762183235867
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, 489.69it/s]
load_f1 = 0.6489795918367347
real_f1 = 0.639511201629328
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, 197.20it/s]
```

1.1.5 Walmart-Amazon - Running the matcher

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

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

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

```
Selected optimization level 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: 0.4475885033607483
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.42828816175460815
step: 20, loss: 0.2503839135169983
step: 30, loss: 0.5751953721046448
step: 40, loss: 0.2362314611673355
step: 50, loss: 0.3353336751461029
step: 60, loss: 0.4737042784690857
step: 70, loss: 0.4466582238674164
step: 80, loss: 0.14500436186790466
step: 90, loss: 0.3159709572792053
step: 100, loss: 0.43465444445610046
step: 110, loss: 0.22180992364883423
step: 120, loss: 0.2927209138870239
step: 130, loss: 0.35581645369529724
step: 140, loss: 0.20999647676944733
step: 150, loss: 0.31380224227905273
step: 160, loss: 0.2404394894838333
step: 170, loss: 0.36342743039131165
step: 180, loss: 0.16178511083126068
step: 190, loss: 0.1826886236667633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.19234642497482374, f1=0.1932220536165908,
best_f1=0.1932220536165908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.366221159696579
step: 10, loss: 0.3097279965877533
step: 20, loss: 0.6155237555503845
step: 30, loss: 0.22069264948368073
step: 40, loss: 0.6505548357963562
step: 50, loss: 0.37513411045074463
step: 60, loss: 0.3564123213291168
step: 70, loss: 0.16036726534366608
step: 80, loss: 0.09192938357591629
step: 90, loss: 0.1503383219242096
step: 100, loss: 0.08930538594722748
step: 110, loss: 0.09435546398162842
step: 120, loss: 0.12556278705596924
step: 130, loss: 0.2694585919380188
step: 140, loss: 0.11449765413999557
step: 150, loss: 0.42743122577667236
step: 160, loss: 0.17965692281723022
step: 170, loss: 0.07773357629776001
step: 180, loss: 0.039742328226566315
step: 190, loss: 0.11938529461622238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7798408488063661, f1=0.7700831024930748,
best_f1=0.7700831024930748
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1967114359140396
step: 10, loss: 0.19436998665332794
step: 20, loss: 0.09095659106969833
step: 30, loss: 0.11319655925035477
step: 40, loss: 0.06467009335756302
step: 50, loss: 0.04993106424808502
step: 60, loss: 0.16807876527309418
step: 70, loss: 0.14116564393043518
step: 80, loss: 0.2015703320503235
```

```
step: 90, loss: 0.09871629625558853
step: 100, loss: 0.05572330579161644
step: 110, loss: 0.35593077540397644
step: 120, loss: 0.3976639211177826
step: 130, loss: 0.23514796793460846
step: 140, loss: 0.031117942184209824
step: 150, loss: 0.24038714170455933
step: 160, loss: 0.1374930441379547
step: 170, loss: 0.21172551810741425
step: 180, loss: 0.11480512470006943
step: 190, loss: 0.12546220421791077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8172043010752689, f1=0.8310249307479225,
best_f1=0.8310249307479225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02615523152053356
step: 10, loss: 0.03874843195080757
step: 20, loss: 0.11315442621707916
step: 30, loss: 0.020373819395899773
step: 40, loss: 0.0663745105266571
step: 50, loss: 0.013963703066110611
step: 60, loss: 0.014147857204079628
step: 70, loss: 0.09795013070106506
step: 80, loss: 0.02640894614160061
step: 90, loss: 0.023074854165315628
step: 100, loss: 0.048751771450042725
step: 110, loss: 0.21436597406864166
step: 120, loss: 0.06762130558490753
step: 130, loss: 0.013565361499786377
step: 140, loss: 0.16174788773059845
step: 150, loss: 0.00780926737934351
step: 160, loss: 0.03217474743723869
step: 170, loss: 0.07442880421876907
step: 180, loss: 0.06841649860143661
step: 190, loss: 0.037077829241752625
/usr/local/lib/python3.7/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=0.8044077134986225, f1=0.8232044198895028,
best_f1=0.8310249307479225
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2186966836452484
step: 10, loss: 0.04158775508403778
step: 20, loss: 0.01259598322212696
step: 30, loss: 0.051576633006334305
step: 40, loss: 0.009757423773407936
step: 50, loss: 0.011780998669564724
step: 60, loss: 0.08177737146615982
step: 70, loss: 0.06229693815112114
step: 80, loss: 0.028980089351534843
step: 90, loss: 0.012718829326331615
step: 100, loss: 0.0450921505689621
step: 110, loss: 0.12430337816476822
step: 120, loss: 0.13239352405071259
step: 130, loss: 0.24195195734500885
step: 140, loss: 0.04591087996959686
step: 150, loss: 0.03499410301446915
step: 160, loss: 0.03064950555562973
step: 170, loss: 0.01763232611119747
step: 180, loss: 0.021019212901592255
step: 190, loss: 0.07922470569610596
/usr/local/lib/python3.7/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=0.8096514745308311, f1=0.8199445983379501,
best_f1=0.8310249307479225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01605825126171112
step: 10, loss: 0.11864925175905228
step: 20, loss: 0.02146838791668415
step: 30, loss: 0.16762515902519226
step: 40, loss: 0.019217517226934433
step: 50, loss: 0.23067991435527802
step: 60, loss: 0.0976928099989891
step: 70, loss: 0.04637369140982628
step: 80, loss: 0.03803076222538948
step: 90, loss: 0.011470364406704903
step: 100, loss: 0.15325133502483368
step: 110, loss: 0.008355554193258286
step: 120, loss: 0.027549095451831818
step: 130, loss: 0.0900956392288208
step: 140, loss: 0.004819451831281185
step: 150, loss: 0.009892567060887814
step: 160, loss: 0.19981934130191803
```

```
step: 170, loss: 0.07032982259988785
step: 180, loss: 0.01538062933832407
step: 190, loss: 0.04862368106842041
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8241758241758241, f1=0.8254847645429362,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015193558298051357
step: 10, loss: 0.007118662353605032
step: 20, loss: 0.0041502052918076515
step: 30, loss: 0.008857849054038525
step: 40, loss: 0.013084116391837597
step: 50, loss: 0.0024798635859042406
step: 60, loss: 0.016054648905992508
step: 70, loss: 0.00678369402885437
step: 80, loss: 0.01617092825472355
step: 90, loss: 0.004401762969791889
step: 100, loss: 0.21258768439292908
step: 110, loss: 0.2014629989862442
step: 120, loss: 0.07654983550310135
step: 130, loss: 0.03524184972047806
step: 140, loss: 0.010423283092677593
step: 150, loss: 0.0073177102021873
step: 160, loss: 0.06318650394678116
step: 170, loss: 0.08990836143493652
step: 180, loss: 0.025152990594506264
step: 190, loss: 0.263904333114624
/usr/local/lib/python3.7/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=0.8142493638676844, f1=0.8041775456919059,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17232848703861237
step: 10, loss: 0.031532980501651764
step: 20, loss: 0.009928696788847446
step: 30, loss: 0.014607928693294525
step: 40, loss: 0.005499434191733599
```

```
step: 50, loss: 0.005621712654829025
step: 60, loss: 0.012220287695527077
step: 70, loss: 0.0017220493173226714
step: 80, loss: 0.19186915457248688
step: 90, loss: 0.028704825788736343
step: 100, loss: 0.0036101429723203182
step: 110, loss: 0.031217394396662712
step: 120, loss: 0.03352510184049606
step: 130, loss: 0.04259948804974556
step: 140, loss: 0.020610105246305466
step: 150, loss: 0.018737087026238844
step: 160, loss: 0.0048723709769546986
step: 170, loss: 0.0021130512468516827
step: 180, loss: 0.010350685566663742
step: 190, loss: 0.011676453053951263
/usr/local/lib/python3.7/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=0.8157894736842105, f1=0.8140161725067385,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003448659786954522
step: 10, loss: 0.0012818946270272136
step: 20, loss: 0.008034836500883102
step: 30, loss: 0.029967816546559334
step: 40, loss: 0.11351336538791656
step: 50, loss: 0.024163508787751198
step: 60, loss: 0.017171215265989304
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.0027628913521766663
step: 80, loss: 0.01173856295645237
step: 90, loss: 0.0025205444544553757
step: 100, loss: 0.02638198435306549
step: 110, loss: 0.0018938021967187524
step: 120, loss: 0.0017657566349953413
step: 130, loss: 0.014888851903378963
step: 140, loss: 0.00577560206875205
step: 150, loss: 0.005106552504003048
step: 160, loss: 0.00590125098824501
step: 170, loss: 0.17867539823055267
step: 180, loss: 0.02254791185259819
step: 190, loss: 0.0015448987251147628
/usr/local/lib/python3.7/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=0.820253164556962, f1=0.8186528497409327,
```

```
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004430172557476908
step: 10, loss: 0.0019524468807503158
step: 20, loss: 0.000606684188824147
step: 30, loss: 0.049846276640892203
step: 40, loss: 0.0012640650384128094
step: 50, loss: 0.017976537346839905
step: 60, loss: 0.0044930847361683846
step: 70, loss: 0.004483567085117102
step: 80, loss: 0.03864945471286774
step: 90, loss: 0.012095210142433643
step: 100, loss: 0.001318741007708013
step: 110, loss: 0.011175939813256264
step: 120, loss: 0.005595720373094082
step: 130, loss: 0.0016307991463690996
step: 140, loss: 0.06328865885734558
step: 150, loss: 0.0005988998454995453
step: 160, loss: 0.12392828613519669
step: 170, loss: 0.0011704294010996819
step: 180, loss: 0.027666263282299042
step: 190, loss: 0.008731084875762463
/usr/local/lib/python3.7/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=0.8219895287958114, f1=0.8247978436657681,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023307574912905693
step: 10, loss: 0.0013191922334954143
step: 20, loss: 0.0025781700387597084
step: 30, loss: 0.028694428503513336
step: 40, loss: 0.0023781785275787115
step: 50, loss: 0.0008788591367192566
step: 60, loss: 0.0029465164989233017
step: 70, loss: 0.0005827071727253497
step: 80, loss: 0.0009293516050092876
step: 90, loss: 0.0010121689410880208
step: 100, loss: 0.005747257731854916
step: 110, loss: 0.170728862285614
step: 120, loss: 0.0018626491073518991
step: 130, loss: 0.03441163897514343
step: 140, loss: 0.014109947718679905
step: 150, loss: 0.015800192952156067
```

```
step: 160, loss: 0.0023683065082877874
step: 170, loss: 0.0024550422094762325
step: 180, loss: 0.0025388780049979687
step: 190, loss: 0.006168984342366457
/usr/local/lib/python3.7/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=0.8113695090439276, f1=0.8095238095238094,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0812058076262474
step: 10, loss: 0.013158811256289482
step: 20, loss: 0.008669768460094929
step: 30, loss: 0.0029285147320479155
step: 40, loss: 0.0018457729602232575
step: 50, loss: 0.0007887452957220376
step: 60, loss: 0.001684939838014543
step: 70, loss: 0.00067613460012353897
step: 80, loss: 0.010061136446893215
step: 90, loss: 0.02156316675245762
step: 100, loss: 0.002398006385192275
step: 110, loss: 0.14456015825271606
step: 120, loss: 0.001199393649585545
step: 130, loss: 0.0003712893812917173
step: 140, loss: 0.0005818352801725268
step: 150, loss: 0.0020605698227882385
step: 160, loss: 0.009114728309214115
step: 170, loss: 0.005835056304931641
step: 180, loss: 0.005294546019285917
step: 190, loss: 0.0011609281646087766
/usr/local/lib/python3.7/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=0.8229426433915212, f1=0.8081841432225065,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23669181764125824
step: 10, loss: 0.0014278244925662875
step: 20, loss: 0.04340382292866707
step: 30, loss: 0.05837707221508026
step: 40, loss: 0.016165366396307945
step: 50, loss: 0.011778759770095348
step: 60, loss: 0.006646502763032913
step: 70, loss: 0.0010316506959497929
```

```
step: 80, loss: 0.01350200641900301
step: 90, loss: 0.0009621192002668977
step: 100, loss: 0.009179637767374516
step: 110, loss: 0.038432229310274124
step: 120, loss: 0.0009419695124961436
step: 130, loss: 0.06293448805809021
step: 140, loss: 0.006810760125517845
step: 150, loss: 0.06829071044921875
step: 160, loss: 0.016795599833130836
step: 170, loss: 0.012559797614812851
step: 180, loss: 0.00450656795874238
step: 190, loss: 0.08514667302370071
/usr/local/lib/python3.7/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=0.8096514745308311, f1=0.8184281842818427,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00699092959985137
step: 10, loss: 0.004708071704953909
step: 20, loss: 0.006202817428857088
step: 30, loss: 0.003221455728635192
step: 40, loss: 0.0039334567263722422
step: 50, loss: 0.013750006444752216
step: 60, loss: 0.039605192840099335
step: 70, loss: 0.0031633020844431052
step: 80, loss: 0.004074579104781151
step: 90, loss: 0.0011954983929172158
step: 100, loss: 0.0020630438812077045
step: 110, loss: 0.00075781531631944656
step: 120, loss: 0.008653220720589161
step: 130, loss: 0.0009313539485447109
step: 140, loss: 0.0011411875020712614
step: 150, loss: 0.00182598817627877
step: 160, loss: 0.002181252231821418
step: 170, loss: 0.0037073406856507063
step: 180, loss: 0.00045176612911745906
step: 190, loss: 0.00817953236401081
/usr/local/lib/python3.7/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=0.8150134048257373, f1=0.8140161725067385,
best_f1=0.8254847645429362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0003817097167484462
step: 10, loss: 0.002316777128726244
step: 20, loss: 0.0036562280729413033
step: 30, loss: 0.0022869817912578583
step: 40, loss: 0.0012700489023700356
step: 50, loss: 0.00111731281504035
step: 60, loss: 0.0018668636912479997
step: 70, loss: 0.0101025914773345
step: 80, loss: 0.0006168699474073946
step: 90, loss: 0.003093633335083723
step: 100, loss: 0.0013956297188997269
step: 110, loss: 0.005215724464505911
step: 120, loss: 0.0021030339412391186
step: 130, loss: 0.002316070720553398
step: 140, loss: 0.0029626740142703056
step: 150, loss: 0.0006840641144663095
step: 160, loss: 0.0014751924900338054
step: 170, loss: 0.027843184769153595
step: 180, loss: 0.0033910972997546196
step: 190, loss: 0.007951609790325165
/usr/local/lib/python3.7/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=0.8138297872340426, f1=0.8184281842818427,
best_f1=0.8254847645429362
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, 240.80it/s]
load_f1 = 0.7989417989417988
real_f1 = 0.7531172069825437
733it [00:00, 3323.53it/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, 195.31it/s]
```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

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

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

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

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

Selected optimization level 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: 0.451819509267807
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5014284253120422
step: 20, loss: 0.32699036598205566
step: 30, loss: 0.4074458181858063
step: 40, loss: 0.5124359130859375
step: 50, loss: 0.30110105872154236

```
step: 60, loss: 0.5737878084182739
step: 70, loss: 0.2999962866306305
step: 80, loss: 0.2897557318210602
step: 90, loss: 0.20344434678554535
step: 100, loss: 0.1753019392490387
step: 110, loss: 0.397659033536911
step: 120, loss: 0.3204011917114258
step: 130, loss: 0.3008592128753662
step: 140, loss: 0.374155730009079
step: 150, loss: 0.29869475960731506
step: 160, loss: 0.39857494831085205
step: 170, loss: 0.3016811013221741
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.22765469824293358, f1=0.24272651704073153,
best_f1=0.24272651704073153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3145139813423157
step: 10, loss: 0.42090439796447754
step: 20, loss: 0.30918824672698975
step: 30, loss: 0.3041732907295227
step: 40, loss: 0.07127591222524643
step: 50, loss: 0.4168875515460968
step: 60, loss: 0.1948815882205963
step: 70, loss: 0.4647762179374695
step: 80, loss: 0.2503117322921753
step: 90, loss: 0.23610612750053406
step: 100, loss: 0.5366449356079102
step: 110, loss: 0.24885790050029755
step: 120, loss: 0.24250416457653046
step: 130, loss: 0.5353748202323914
step: 140, loss: 0.52547687292099
step: 150, loss: 0.4243218004703522
step: 160, loss: 0.4164438545703888
step: 170, loss: 0.39727485179901123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.24272651704073153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6099634766578674
step: 10, loss: 0.32928740978240967
step: 20, loss: 0.2353951632976532
step: 30, loss: 0.24676433205604553
step: 40, loss: 0.36263349652290344
step: 50, loss: 0.6343328952789307
step: 60, loss: 0.3095738887786865
step: 70, loss: 0.2382466197013855
step: 80, loss: 0.3691555857658386
step: 90, loss: 0.5629407167434692
step: 100, loss: 0.2999214828014374
step: 110, loss: 0.11915397644042969
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.5667478442192078
step: 130, loss: 0.5166899561882019
step: 140, loss: 0.4417756199836731
step: 150, loss: 0.19408182799816132
step: 160, loss: 0.1619350016117096
step: 170, loss: 0.28876763582229614
/usr/local/lib/python3.7/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=0.20189905047476264, f1=0.2058074375955171,
best_f1=0.24272651704073153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3616172969341278
step: 10, loss: 0.511970043182373
step: 20, loss: 0.21758893132209778
step: 30, loss: 0.40882623195648193
step: 40, loss: 0.23120051622390747
step: 50, loss: 0.32430723309516907
step: 60, loss: 0.6099957823753357
step: 70, loss: 0.30672362446784973
step: 80, loss: 0.48294374346733093
step: 90, loss: 0.28634971380233765
step: 100, loss: 0.3543798327445984
step: 110, loss: 0.44221216440200806
step: 120, loss: 0.5809059739112854
step: 130, loss: 0.2350456714630127
step: 140, loss: 0.28520166873931885
step: 150, loss: 0.5486828684806824
step: 160, loss: 0.10564465820789337
step: 170, loss: 0.2563264071941376
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3795782463928967, f1=0.36184210526315785,
best_f1=0.36184210526315785
/usr/local/lib/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 4096.0
step: 0, loss: 0.3541887402534485
step: 10, loss: 0.3391483724117279
step: 20, loss: 0.30963560938835144
step: 30, loss: 0.25446027517318726
step: 40, loss: 0.23840467631816864
step: 50, loss: 0.16111968457698822
step: 60, loss: 0.30474787950515747
step: 70, loss: 0.2681099474430084
step: 80, loss: 0.10826854407787323
step: 90, loss: 0.5002583861351013
step: 100, loss: 0.17261572182178497
step: 110, loss: 0.304511159658432
step: 120, loss: 0.10745584964752197
step: 130, loss: 0.2318611443042755
step: 140, loss: 0.14839120209217072
step: 150, loss: 0.3482333719730377
step: 160, loss: 0.14507541060447693
step: 170, loss: 0.14158634841442108
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5806451612903226, f1=0.5764023210831721,
best_f1=0.5764023210831721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26858144998550415
step: 10, loss: 0.10366956144571304
step: 20, loss: 0.3117378354072571
step: 30, loss: 0.10806263983249664
step: 40, loss: 0.16048207879066467
step: 50, loss: 0.22676974534988403
```

```
step: 60, loss: 0.10152298212051392
step: 70, loss: 0.2957876920700073
step: 80, loss: 0.1030137687921524
step: 90, loss: 0.32992011308670044
step: 100, loss: 0.1177893877029419
step: 110, loss: 0.33631768822669983
step: 120, loss: 0.2551504075527191
step: 130, loss: 0.3931455612182617
step: 140, loss: 0.14758406579494476
step: 150, loss: 0.22364971041679382
step: 160, loss: 0.25274282693862915
step: 170, loss: 0.1881512999534607
/usr/local/lib/python3.7/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=0.5463576158940397, f1=0.5537918871252204,
best_f1=0.5764023210831721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20053212344646454
step: 10, loss: 0.43572917580604553
step: 20, loss: 0.41722846031188965
step: 30, loss: 0.270607590675354
step: 40, loss: 0.0818043202161789
step: 50, loss: 0.1548708975315094
step: 60, loss: 0.5859184265136719
step: 70, loss: 0.30775186419487
step: 80, loss: 0.10672281682491302
step: 90, loss: 0.2545713484287262
step: 100, loss: 0.18856948614120483
step: 110, loss: 0.25070345401763916
step: 120, loss: 0.3627302646636963
step: 130, loss: 0.07284790277481079
step: 140, loss: 0.14603914320468903
step: 150, loss: 0.20420916378498077
step: 160, loss: 0.09259164333343506
step: 170, loss: 0.11691600829362869
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5811320754716981, f1=0.6034816247582204,
best_f1=0.6034816247582204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15021218359470367
step: 10, loss: 0.17764022946357727
step: 20, loss: 0.14452095329761505
step: 30, loss: 0.09409407526254654
step: 40, loss: 0.14646311104297638
step: 50, loss: 0.08885465562343597
step: 60, loss: 0.19850611686706543
step: 70, loss: 0.1676354557275772
step: 80, loss: 0.1663978546857834
step: 90, loss: 0.16353705525398254
step: 100, loss: 0.11971999704837799
step: 110, loss: 0.3726990520954132
step: 120, loss: 0.3876909911632538
step: 130, loss: 0.18853211402893066
step: 140, loss: 0.2380373179912567
step: 150, loss: 0.16860927641391754
step: 160, loss: 0.09543155878782272
step: 170, loss: 0.3156023919582367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6439232409381663, f1=0.6371308016877637,
best_f1=0.6371308016877637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16954688727855682
step: 10, loss: 0.1458640843629837
step: 20, loss: 0.16725443303585052
step: 30, loss: 0.12040664255619049
step: 40, loss: 0.32068267464637756
step: 50, loss: 0.047931402921676636
step: 60, loss: 0.06496334075927734
step: 70, loss: 0.36791977286338806
step: 80, loss: 0.1130114495754242
step: 90, loss: 0.14882051944732666
step: 100, loss: 0.163815438747406
step: 110, loss: 0.16520839929580688
step: 120, loss: 0.34300392866134644
step: 130, loss: 0.20315833389759064
step: 140, loss: 0.12170226126909256
step: 150, loss: 0.3496345579624176
```

```
step: 160, loss: 0.07425655424594879
step: 170, loss: 0.05834255740046501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.66260162601626, f1=0.6611909650924026,
best_f1=0.6611909650924026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17304201424121857
step: 10, loss: 0.17026454210281372
step: 20, loss: 0.14416880905628204
step: 30, loss: 0.2595047652721405
step: 40, loss: 0.15476645529270172
step: 50, loss: 0.11696650832891464
step: 60, loss: 0.1866922378540039
step: 70, loss: 0.21492646634578705
step: 80, loss: 0.12042732536792755
step: 90, loss: 0.09900413453578949
step: 100, loss: 0.17166782915592194
step: 110, loss: 0.08870799839496613
step: 120, loss: 0.09254252910614014
step: 130, loss: 0.1388843059539795
step: 140, loss: 0.11203150451183319
step: 150, loss: 0.1833127737045288
step: 160, loss: 0.1522139608860016
step: 170, loss: 0.14059679210186005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7058823529411765, f1=0.6724511930585683,
best_f1=0.6724511930585683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2978176176548004
step: 10, loss: 0.12591587007045746
step: 20, loss: 0.019166098907589912
step: 30, loss: 0.4635078012943268
```

```
step: 40, loss: 0.054665692150592804
step: 50, loss: 0.035297878086566925
step: 60, loss: 0.1675274819135666
step: 70, loss: 0.07855430990457535
step: 80, loss: 0.11044832319021225
step: 90, loss: 0.07631973177194595
step: 100, loss: 0.3032678961753845
step: 110, loss: 0.12281856685876846
step: 120, loss: 0.12887148559093475
step: 130, loss: 0.11053312569856644
step: 140, loss: 0.21984131634235382
step: 150, loss: 0.056399956345558167
step: 160, loss: 0.11215739697217941
step: 170, loss: 0.2633460760116577
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7325301204819279, f1=0.7061611374407583,
best_f1=0.7061611374407583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10130393505096436
step: 10, loss: 0.11668038368225098
step: 20, loss: 0.09328115731477737
step: 30, loss: 0.17712852358818054
step: 40, loss: 0.02959374710917473
step: 50, loss: 0.1500551700592041
step: 60, loss: 0.17865461111068726
step: 70, loss: 0.057447534054517746
step: 80, loss: 0.05705129727721214
step: 90, loss: 0.0441376231610775
step: 100, loss: 0.06319783627986908
step: 110, loss: 0.1123826801776886
step: 120, loss: 0.04348526895046234
step: 130, loss: 0.17625072598457336
step: 140, loss: 0.1018265038728714
step: 150, loss: 0.06008574366569519
step: 160, loss: 0.23603664338588715
step: 170, loss: 0.27152693271636963
/usr/local/lib/python3.7/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=0.7289293849658314, f1=0.7058823529411764,
```

```
best_f1=0.7061611374407583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11586257815361023
step: 10, loss: 0.038828227669000626
step: 20, loss: 0.08019572496414185
step: 30, loss: 0.15752221643924713
step: 40, loss: 0.1264435350894928
step: 50, loss: 0.04709884524345398
step: 60, loss: 0.11469627171754837
step: 70, loss: 0.19999772310256958
step: 80, loss: 0.02574087493121624
step: 90, loss: 0.08027614653110504
step: 100, loss: 0.18802788853645325
step: 110, loss: 0.08121108263731003
step: 120, loss: 0.05921945348381996
step: 130, loss: 0.0520111583173275
step: 140, loss: 0.11512961238622665
step: 150, loss: 0.06918235123157501
step: 160, loss: 0.13063186407089233
step: 170, loss: 0.03701460734009743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7389162561576353, f1=0.7320574162679426,
best_f1=0.7320574162679426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09354796260595322
step: 10, loss: 0.11353924125432968
step: 20, loss: 0.08058355748653412
step: 30, loss: 0.09878221154212952
step: 40, loss: 0.10853340476751328
step: 50, loss: 0.0457746796309948
step: 60, loss: 0.12391617149114609
step: 70, loss: 0.10579171776771545
step: 80, loss: 0.20667129755020142
step: 90, loss: 0.0274376031011343
step: 100, loss: 0.14895719289779663
step: 110, loss: 0.23302146792411804
step: 120, loss: 0.057980652898550034
step: 130, loss: 0.2248634248971939
```

```
step: 140, loss: 0.04676015302538872
step: 150, loss: 0.02273234911262989
step: 160, loss: 0.0862942561507225
step: 170, loss: 0.10026071220636368
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7434052757793764, f1=0.7393364928909952,
best_f1=0.7393364928909952
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013402208685874939
step: 10, loss: 0.0127119654789567
step: 20, loss: 0.029604125767946243
step: 30, loss: 0.03917897120118141
step: 40, loss: 0.010154898278415203
step: 50, loss: 0.06471972167491913
step: 60, loss: 0.024232447147369385
step: 70, loss: 0.11450003832578659
step: 80, loss: 0.1316300928592682
step: 90, loss: 0.1109597235918045
step: 100, loss: 0.2159043550491333
step: 110, loss: 0.11901721358299255
step: 120, loss: 0.04064543917775154
step: 130, loss: 0.08035513013601303
step: 140, loss: 0.07933130115270615
step: 150, loss: 0.04445789009332657
step: 160, loss: 0.041316766291856766
step: 170, loss: 0.02846718207001686
/usr/local/lib/python3.7/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=0.7418546365914787, f1=0.7493917274939174,
best_f1=0.7393364928909952
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
```

```
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:05, 322.36it/s]
load_f1 = 0.5608856088560885
real_f1 = 0.5522388059701493
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, 196.40it/s]
```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

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

```
  --save_model \
  --da entry_swap  \
  --da span_shuffle \
  --summarize

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 413kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 804kB/s]
Downloading: 100% 456k/456k [00:01<00:00, 399kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 64.8MB/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: 0.5597564578056335
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4200647473335266
step: 20, loss: 0.52981615066652832
step: 30, loss: 0.3215363025665283
step: 40, loss: 0.3178674280643463
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.252368688583374
step: 60, loss: 0.1536213606595993
step: 70, loss: 0.03558827564120293
step: 80, loss: 0.01784016191959381
step: 90, loss: 0.20114505290985107
step: 100, loss: 0.0323970764875412
step: 110, loss: 0.08986739814281464
step: 120, loss: 0.014277835376560688
step: 130, loss: 0.025940634310245514
step: 140, loss: 0.11856900155544281
step: 150, loss: 0.16757549345493317
step: 160, loss: 0.10336750000715256
step: 170, loss: 0.038880862295627594
step: 180, loss: 0.1918790489435196
step: 190, loss: 0.0280488058924675
step: 200, loss: 0.10995873063802719
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 210, loss: 0.1479366570711136
step: 220, loss: 0.12099909782409668
step: 230, loss: 0.009206313639879227
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9630459126539753, f1=0.9617977528089887,
best_f1=0.9617977528089887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005314359907060862
step: 10, loss: 0.3042723834514618
step: 20, loss: 0.01546530518680811
step: 30, loss: 0.05040949583053589
step: 40, loss: 0.06875249743461609
step: 50, loss: 0.005910285282880068
step: 60, loss: 0.04595309495925903
step: 70, loss: 0.0031606554985046387
step: 80, loss: 0.0022513028234243393
step: 90, loss: 0.01861126348376274
step: 100, loss: 0.019218783825663591
step: 110, loss: 0.08641225844621658
step: 120, loss: 0.024518977850675583
step: 130, loss: 0.020003657788038254
step: 140, loss: 0.0036132631357759237
step: 150, loss: 0.030064821243286133
step: 160, loss: 0.011748861521482468
step: 170, loss: 0.013651725836098194
step: 180, loss: 0.011066245846450329
step: 190, loss: 0.03854929283261299
step: 200, loss: 0.0118034016340971
step: 210, loss: 0.09820117801427841
step: 220, loss: 0.017934665083885193
step: 230, loss: 0.0023978909011930227
/usr/local/lib/python3.7/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=0.9515418502202643, f1=0.9551569506726457,
best_f1=0.9617977528089887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05550415441393852
step: 10, loss: 0.020417384803295135
step: 20, loss: 0.052510444074869156
step: 30, loss: 0.03708773851394653
step: 40, loss: 0.0028005014173686504
step: 50, loss: 0.08011089265346527
step: 60, loss: 0.04799386113882065
step: 70, loss: 0.01766725443303585
step: 80, loss: 0.1005350723862648
step: 90, loss: 0.041998494416475296
step: 100, loss: 0.0029734899289906025
step: 110, loss: 0.002418916206806898
step: 120, loss: 0.0008030786411836743
step: 130, loss: 0.12409982830286026
```

```
step: 140, loss: 0.0072603989392519
step: 150, loss: 0.02634747512638569
step: 160, loss: 0.020444342866539955
step: 170, loss: 0.0070685516111552715
step: 180, loss: 0.1652350276708603
step: 190, loss: 0.02665121853351593
step: 200, loss: 0.007569437846541405
step: 210, loss: 0.005104101728647947
step: 220, loss: 0.08270000666379929
step: 230, loss: 0.010018005035817623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9683972911963882, f1=0.9651293588301463,
best_f1=0.9651293588301463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03351811319589615
step: 10, loss: 0.0025437711738049984
step: 20, loss: 0.006027617026120424
step: 30, loss: 0.0013345615006983828
step: 40, loss: 0.05089115723967552
step: 50, loss: 0.012857252731919289
step: 60, loss: 0.009150496684014797
step: 70, loss: 0.17789874970912933
step: 80, loss: 0.1336848884820938
step: 90, loss: 0.06590500473976135
step: 100, loss: 0.012555013410747051
step: 110, loss: 0.003292706096544862
step: 120, loss: 0.008303779177367687
step: 130, loss: 0.007318594958633184
step: 140, loss: 0.018647821620106697
step: 150, loss: 0.026101497933268547
step: 160, loss: 0.007846538908779621
step: 170, loss: 0.0022930533159524202
step: 180, loss: 0.08199869096279144
step: 190, loss: 0.002942116931080818
step: 200, loss: 0.06647658348083496
step: 210, loss: 0.034762442111968994
step: 220, loss: 0.0016831576358526945
step: 230, loss: 0.036640096455812454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9732142857142857, f1=0.9694224235560589,
best_f1=0.9694224235560589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012140420731157064
step: 10, loss: 0.00192552024964224317
step: 20, loss: 0.0013578888028860092
step: 30, loss: 0.0012194678420200944
step: 40, loss: 0.007247534580528736
step: 50, loss: 0.0012849910417571664
step: 60, loss: 0.02902986854314804
step: 70, loss: 0.011176178231835365
step: 80, loss: 0.19974835216999054
step: 90, loss: 0.2432413101196289
step: 100, loss: 0.016194382682442665
step: 110, loss: 0.0021194329019635916
step: 120, loss: 0.0035481436643749475
step: 130, loss: 0.0017605902394279838
step: 140, loss: 0.009316902607679367
step: 150, loss: 0.002824636874720454
step: 160, loss: 0.001963953021913767
step: 170, loss: 0.02802020125091076
step: 180, loss: 0.0015926172491163015
step: 190, loss: 0.016639452427625656
step: 200, loss: 0.009152630344033241
step: 210, loss: 0.005777392070740461
step: 220, loss: 0.0015007101465016603
step: 230, loss: 0.02516115829348564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775784753363228, f1=0.9774266365688488,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005760454805567861
step: 10, loss: 0.0036884844303131104
```

```
step: 20, loss: 0.030140740796923637
step: 30, loss: 0.0005238145822659135
step: 40, loss: 0.00024504499742761254
step: 50, loss: 0.0005645704222843051
step: 60, loss: 0.0009482656605541706
step: 70, loss: 0.009287613444030285
step: 80, loss: 0.008770667016506195
step: 90, loss: 0.03269057720899582
step: 100, loss: 0.004596272483468056
step: 110, loss: 0.025469956919550896
step: 120, loss: 0.002942050574347377
step: 130, loss: 0.001915855216793716
step: 140, loss: 0.0016250902554020286
step: 150, loss: 0.000596984347794205
step: 160, loss: 0.013436437584459782
step: 170, loss: 0.0005704913637600839
step: 180, loss: 0.001494323369115591
step: 190, loss: 0.00034095608862116933
step: 200, loss: 0.0015699259238317609
step: 210, loss: 0.0025788494385778904
step: 220, loss: 0.03617382422089577
step: 230, loss: 0.004934875760227442
/usr/local/lib/python3.7/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=0.9765363128491621, f1=0.9673790776152981,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005092521198093891
step: 10, loss: 0.0008612816454842687
step: 20, loss: 0.0019105918472632766
step: 30, loss: 0.00045340292854234576
step: 40, loss: 0.001323989941738546
step: 50, loss: 0.0009882488520815969
step: 60, loss: 0.0005044252029620111
step: 70, loss: 0.0005792967276647687
step: 80, loss: 0.0007343297475017607
step: 90, loss: 0.002223563613370061
step: 100, loss: 0.00445591239258647
step: 110, loss: 0.0006574684521183372
step: 120, loss: 0.0022370500955730677
step: 130, loss: 0.0017217834247276187
step: 140, loss: 0.0006241808878257871
step: 150, loss: 0.0038042166270315647
step: 160, loss: 0.00045605076593346894
step: 170, loss: 0.0013419000897556543
```

```
step: 180, loss: 0.0030220302287489176
step: 190, loss: 0.004147415980696678
step: 200, loss: 0.00255036074668169
step: 210, loss: 0.05815616622567177
step: 220, loss: 0.0023459780495613813
step: 230, loss: 0.0021406987216323614
/usr/local/lib/python3.7/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=0.9755555555555556, f1=0.9721913236929923,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003414815291762352
step: 10, loss: 0.024609185755252838
step: 20, loss: 0.0018466792535036802
step: 30, loss: 0.001169582479633391
step: 40, loss: 0.0025680074468255043
step: 50, loss: 0.0011203784961253405
step: 60, loss: 0.0008967284811660647
step: 70, loss: 0.0003407771873753518
step: 80, loss: 0.026320308446884155
step: 90, loss: 0.00025796532281674445
step: 100, loss: 0.00035405418020673096
step: 110, loss: 0.0006555730360560119
step: 120, loss: 0.0005947107565589249
step: 130, loss: 0.0005563297891058028
step: 140, loss: 0.0002697837189771235
step: 150, loss: 0.26286646723747253
step: 160, loss: 0.0004561288224067539
step: 170, loss: 0.0655713900923729
step: 180, loss: 0.00045890803448855877
step: 190, loss: 0.0006290409946814179
step: 200, loss: 0.03264978528022766
step: 210, loss: 0.0019451961852610111
step: 220, loss: 0.0025045534130185843
step: 230, loss: 0.0006082129548303783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776785714285714, f1=0.9788182831661093,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008929232135415077
step: 10, loss: 0.0006474928231909871
step: 20, loss: 0.0010529608698561788
step: 30, loss: 0.0017688819207251072
step: 40, loss: 0.002984401071444154
step: 50, loss: 0.0017905380809679627
step: 60, loss: 0.000683360907714814
step: 70, loss: 0.0022888665553182364
step: 80, loss: 0.0002559188869781792
step: 90, loss: 0.18704837560653687
step: 100, loss: 0.0022636130452156067
step: 110, loss: 0.0008178668795153499
step: 120, loss: 0.0018289211438968778
step: 130, loss: 0.0007880459888838232
step: 140, loss: 0.0006595907616429031
step: 150, loss: 0.0017509317258372903
step: 160, loss: 0.039281345903873444
step: 170, loss: 0.000284697423921898
step: 180, loss: 0.000896357640158385
step: 190, loss: 0.03730888292193413
step: 200, loss: 0.000471060979180038
step: 210, loss: 0.0009786098962649703
step: 220, loss: 0.0004701297730207443
step: 230, loss: 0.000550552736967802
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787709497206705, f1=0.9774774774774775,
best_f1=0.9774774774774775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00034876557765552856
step: 10, loss: 0.0005284378421492875
step: 20, loss: 0.0019661851692944765
step: 30, loss: 0.00030644042999483645
step: 40, loss: 0.0046006496995687485
step: 50, loss: 0.000278875813819468
step: 60, loss: 0.0005244702333584428
step: 70, loss: 0.0055853514932096004
step: 80, loss: 0.015821034088730812
step: 90, loss: 0.0005576180992648005
```

```
step: 100, loss: 0.00026491325115785
step: 110, loss: 0.003924877382814884
step: 120, loss: 0.0002752003201749176
step: 130, loss: 0.0002622239990159869
step: 140, loss: 0.00019217714725527912
step: 150, loss: 0.0028589500579982996
step: 160, loss: 9.245253750123084e-05
step: 170, loss: 0.00012565701035782695
step: 180, loss: 0.0002607532369438559
step: 190, loss: 0.02981644868850708
step: 200, loss: 0.0001964702532859519
step: 210, loss: 0.003185007721185684
step: 220, loss: 0.001356669352389872
step: 230, loss: 0.0002148685889551416
/usr/local/lib/python3.7/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=0.9711751662971175, f1=0.9688888888888889,
best_f1=0.9774774774774775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003024709876662673
step: 10, loss: 0.023698972538113594
step: 20, loss: 0.00015654900926165283
step: 30, loss: 0.00031020346796140075
step: 40, loss: 3.5781675251200795e-05
step: 50, loss: 0.0002782711526378989
step: 60, loss: 0.017723100259900093
step: 70, loss: 0.00017674190166871995
step: 80, loss: 0.00019807468925137073
step: 90, loss: 0.17048169672489166
step: 100, loss: 0.0005980055429972708
step: 110, loss: 0.0006197079783305526
step: 120, loss: 0.00024416937958449125
step: 130, loss: 0.00042040564585477114
step: 140, loss: 0.0008361805812455714
step: 150, loss: 0.0005038549425080419
step: 160, loss: 0.0032541591208428144
step: 170, loss: 0.0013236927334219217
step: 180, loss: 0.0007429346442222595
step: 190, loss: 0.0004353647236712277
step: 200, loss: 0.006023638881742954
step: 210, loss: 0.0005980512360110879
step: 220, loss: 0.0025820014998316765
step: 230, loss: 0.0006854443927295506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.9796839729119639,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006562226917594671
step: 10, loss: 0.00027973006945103407
step: 20, loss: 0.0005263894563540816
step: 30, loss: 0.004435927607119083
step: 40, loss: 0.0002495044900570065
step: 50, loss: 0.000935097225010395
step: 60, loss: 0.00015640912170056254
step: 70, loss: 0.00011330361303407699
step: 80, loss: 5.6049095292109996e-05
step: 90, loss: 0.009315119124948978
step: 100, loss: 7.281794387381524e-05
step: 110, loss: 5.999468339723535e-05
step: 120, loss: 0.001233310904353857
step: 130, loss: 0.00016725435853004456
step: 140, loss: 0.000292012031422928
step: 150, loss: 0.00018164263747166842
step: 160, loss: 0.002177042653784156
step: 170, loss: 0.00015899515710771084
step: 180, loss: 8.751273708185181e-05
step: 190, loss: 0.0011120050912722945
step: 200, loss: 0.00010466889943927526
step: 210, loss: 0.00010813036351464689
step: 220, loss: 0.0013768752105534077
step: 230, loss: 0.00036253733560442924
/usr/local/lib/python3.7/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=0.9798206278026906, f1=0.9762174405436014,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035200282000005245
step: 10, loss: 0.00010960047802655026
step: 20, loss: 0.00014828004350420088
step: 30, loss: 0.0005736736347898841
step: 40, loss: 0.00019069014524575323
step: 50, loss: 0.002974599367007613
```

```
step: 60, loss: 0.00012216750474181026
step: 70, loss: 0.0005657640285789967
step: 80, loss: 0.000223309631110169
step: 90, loss: 0.00013771865633316338
step: 100, loss: 0.0008180536096915603
step: 110, loss: 0.00032839624327607453
step: 120, loss: 0.00014841977099422365
step: 130, loss: 0.00017265793576370925
step: 140, loss: 8.19780834717676e-05
step: 150, loss: 8.64747999003157e-05
step: 160, loss: 0.002788471058011055
step: 170, loss: 0.0002550910576246679
step: 180, loss: 0.0017581578576937318
step: 190, loss: 0.0001597844820935279
step: 200, loss: 5.852121830685064e-05
step: 210, loss: 0.00605862308293581
step: 220, loss: 0.0001784753694664687
step: 230, loss: 0.00011343045480316505
/usr/local/lib/python3.7/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=0.9774774774774775, f1=0.9727272727272728,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.820369890192524e-05
step: 10, loss: 7.614267815370113e-05
step: 20, loss: 0.0009846310131251812
step: 30, loss: 0.00012651423458009958
step: 40, loss: 0.00013301237777341157
step: 50, loss: 7.914703019196168e-05
step: 60, loss: 0.00037233831244520843
step: 70, loss: 0.0004165015707258135
step: 80, loss: 0.00021451186330523342
step: 90, loss: 0.00016980546934064478
step: 100, loss: 0.00014417622878681868
step: 110, loss: 0.004817027132958174
step: 120, loss: 2.9405080567812547e-05
step: 130, loss: 0.00015712712774984539
step: 140, loss: 8.90814044396393e-05
step: 150, loss: 6.140084587968886e-05
step: 160, loss: 0.0002647697110660374
step: 170, loss: 0.00011829427967313677
step: 180, loss: 7.033632573438808e-05
step: 190, loss: 0.000129058986203745
step: 200, loss: 0.0005928740138188004
step: 210, loss: 0.00024251166905742139
```

```
step: 220, loss: 8.639995940029621e-05
step: 230, loss: 0.0002614410186652094
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9739524348810873,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014895277854520828
step: 10, loss: 8.124222949845716e-05
step: 20, loss: 0.0001374067651340738
step: 30, loss: 6.558827590197325e-05
step: 40, loss: 4.131040259380825e-05
step: 50, loss: 5.227232031757012e-05
step: 60, loss: 0.000202896015252918
step: 70, loss: 0.00022837224241811782
step: 80, loss: 5.8689634897746146e-05
step: 90, loss: 4.871827331953682e-05
step: 100, loss: 4.6084558562142774e-05
step: 110, loss: 6.827122706454247e-05
step: 120, loss: 0.08321920037269592
step: 130, loss: 5.620247247861698e-05
step: 140, loss: 0.0002171196392737329
step: 150, loss: 7.611831824760884e-05
step: 160, loss: 0.000473716645501554
step: 170, loss: 5.9878992033191025e-05
step: 180, loss: 7.2106508014258e-05
step: 190, loss: 0.00020319217583164573
step: 200, loss: 0.00023083161795511842
step: 210, loss: 0.003593367524445057
step: 220, loss: 6.336069054668769e-05
step: 230, loss: 9.08168003661558e-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=0.9797752808988766, f1=0.9751131221719457,
best_f1=0.9796839729119639
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, 155.34it/s]
load_f1 = 0.9766407119021134
real_f1 = 0.9755555555555556
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:31, 139.29it/s]
```

1.3.2 DBLP-GoogleScholar - Running the matcher

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

```
  --da attr_del \
  --summarize

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

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

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

Defaults for this optimization level are:
```
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
```
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
```
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
```
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6224334836006165
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.45402300357818604
step: 20, loss: 0.3502412736415863

```
step: 30, loss: 0.37261006236076355
step: 40, loss: 0.19683898985385895
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.17826327681541443
step: 60, loss: 0.18448826670646667
step: 70, loss: 0.14887690544128418
step: 80, loss: 0.09249886870384216
step: 90, loss: 0.1920507252216339
step: 100, loss: 0.22116932272911072
step: 110, loss: 0.11910207569599152
step: 120, loss: 0.06445236504077911
step: 130, loss: 0.06605061143636703
step: 140, loss: 0.2810341417789459
step: 150, loss: 0.05704644322395325
step: 160, loss: 0.2276923954486847
step: 170, loss: 0.04840584099292755
step: 180, loss: 0.2205246239900589
step: 190, loss: 0.08571866899728775
step: 200, loss: 0.0672835186123848
step: 210, loss: 0.0812624990940094
step: 220, loss: 0.09241444617509842
step: 230, loss: 0.21231548488140106
step: 240, loss: 0.0817226991057396
step: 250, loss: 0.03862329572439194
step: 260, loss: 0.205467090010643
step: 270, loss: 0.30667275190353394
step: 280, loss: 0.03424863889813423
step: 290, loss: 0.07033836096525192
step: 300, loss: 0.10571275651454926
step: 310, loss: 0.2244005650281906
step: 320, loss: 0.089214026927948
step: 330, loss: 0.09762924909591675
step: 340, loss: 0.30943915247917175
step: 350, loss: 0.09077880531549454
step: 360, loss: 0.04236092418432236
step: 370, loss: 0.07794227451086044
step: 380, loss: 0.09837373346090317
step: 390, loss: 0.0186260174959898
step: 400, loss: 0.0583733581006526950
step: 410, loss: 0.24065028131008148
step: 420, loss: 0.05329645797610283
step: 430, loss: 0.050691910088062286
step: 440, loss: 0.06160971149802208
step: 450, loss: 0.010186858475208282
step: 460, loss: 0.019831962883472443
step: 470, loss: 0.06366570293903351
step: 480, loss: 0.14044708013534546
step: 490, loss: 0.089410699903965
```

```
step: 500, loss: 0.03312423452734947
step: 510, loss: 0.0521097332239151
step: 520, loss: 0.2510397136211395
step: 530, loss: 0.04342237859964371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9354078264969354, f1=0.9321478708469817,
best_f1=0.9321478708469817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01931505650281906
step: 10, loss: 0.026738125830888748
step: 20, loss: 0.019148560240864754
step: 30, loss: 0.020868409425020218
step: 40, loss: 0.08579752594232559
step: 50, loss: 0.0641045942902565
step: 60, loss: 0.14911919832229614
step: 70, loss: 0.039802320301532745
step: 80, loss: 0.0456983707845211
step: 90, loss: 0.03745950758457184
step: 100, loss: 0.15929022431373596
step: 110, loss: 0.018375009298324585
step: 120, loss: 0.0938703641295433
step: 130, loss: 0.032649390399456024
step: 140, loss: 0.013526102527976036
step: 150, loss: 0.052014704793691635
step: 160, loss: 0.020746314898133278
step: 170, loss: 0.03480425849556923
step: 180, loss: 0.04185085371136665
step: 190, loss: 0.004960999358445406
step: 200, loss: 0.2768349051475525
step: 210, loss: 0.035925351083278656
step: 220, loss: 0.001765235560014844
step: 230, loss: 0.029451362788677216
step: 240, loss: 0.15757916867733002
step: 250, loss: 0.01687079295516014
step: 260, loss: 0.088102787733078
step: 270, loss: 0.07692565768957138
step: 280, loss: 0.1628410965204239
step: 290, loss: 0.031007083132863045
step: 300, loss: 0.0477481484413147
step: 310, loss: 0.0385958030819892
```

```
step: 320, loss: 0.07601764798164368
step: 330, loss: 0.0766957700252533
step: 340, loss: 0.07110867649316788
step: 350, loss: 0.004245857708156109
step: 360, loss: 0.11141841113567352
step: 370, loss: 0.00953134149312973
step: 380, loss: 0.1591448336839676
step: 390, loss: 0.005387195385992527
step: 400, loss: 0.09505664557218552
step: 410, loss: 0.03662283346056938
step: 420, loss: 0.0888509526848793
step: 430, loss: 0.11071154475212097
step: 440, loss: 0.015582920052111149
step: 450, loss: 0.05173176899552345
step: 460, loss: 0.06956041604280472
step: 470, loss: 0.03994572162628174
step: 480, loss: 0.12731577455997467
step: 490, loss: 0.08818350732326508
step: 500, loss: 0.01312828715890646
step: 510, loss: 0.03563634678721428
step: 520, loss: 0.3929632008075714
step: 530, loss: 0.1755896657705307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9360331339162449, f1=0.9394495412844037,
best_f1=0.9394495412844037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09061792492866516
step: 10, loss: 0.039090629667043686
step: 20, loss: 0.05549672618508339
step: 30, loss: 0.07158049196004868
step: 40, loss: 0.06898995488882065
step: 50, loss: 0.017505913972854614
step: 60, loss: 0.022725753486156464
step: 70, loss: 0.010083588771522045
step: 80, loss: 0.02515694871544838
step: 90, loss: 0.006876911967992783
step: 100, loss: 0.045436691492795944
step: 110, loss: 0.09146624058485031
step: 120, loss: 0.06680289655923843
step: 130, loss: 0.014804719015955925
```

```
step: 140, loss: 0.1213110014796257
step: 150, loss: 0.02892627753317356
step: 160, loss: 0.09851682931184769
step: 170, loss: 0.05712664872407913
step: 180, loss: 0.022823531180620193
step: 190, loss: 0.044076740741729736
step: 200, loss: 0.022078845649957657
step: 210, loss: 0.06893593072891235
step: 220, loss: 0.1400143951177597
step: 230, loss: 0.08213300257921219
step: 240, loss: 0.023541854694485664
step: 250, loss: 0.17236755788326263
step: 260, loss: 0.17551562190055847
step: 270, loss: 0.07454731315374374
step: 280, loss: 0.003859067801386118
step: 290, loss: 0.011866024695336819
step: 300, loss: 0.07172142714262009
step: 310, loss: 0.0575706847012043
step: 320, loss: 0.034134913235902786
step: 330, loss: 0.038777049630880356
step: 340, loss: 0.023973343893885612
step: 350, loss: 0.17162449657917023
step: 360, loss: 0.006659659091383219
step: 370, loss: 0.04862646013498306
step: 380, loss: 0.05343804880976677
step: 390, loss: 0.024585075676441193
step: 400, loss: 0.033497367054224014
step: 410, loss: 0.08471423387527466
step: 420, loss: 0.010953215882182121
step: 430, loss: 0.042358819395303726
step: 440, loss: 0.27081167697906494
step: 450, loss: 0.07279501110315323
step: 460, loss: 0.12826375663280487
step: 470, loss: 0.007268426939845085
step: 480, loss: 0.05292513966560364
step: 490, loss: 0.032530371099710464
step: 500, loss: 0.02269849181175232
step: 510, loss: 0.01748614013195038
step: 520, loss: 0.0044626700691878796
step: 530, loss: 0.014065772294998169
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9416126042632066, f1=0.9454545454545454,
```

```
best_f1=0.9454545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007792674470692873
step: 10, loss: 0.0038558579981327057
step: 20, loss: 0.12855522334575653
step: 30, loss: 0.147822767496109
step: 40, loss: 0.012929561547935009
step: 50, loss: 0.05262710526585579
step: 60, loss: 0.03902212157845497
step: 70, loss: 0.08791710436344147
step: 80, loss: 0.15544331073760986
step: 90, loss: 0.014380390755832195
step: 100, loss: 0.002300231484696269
step: 110, loss: 0.15998244285583496
step: 120, loss: 0.0035437624901533127
step: 130, loss: 0.18223918974399567
step: 140, loss: 0.03920836001634598
step: 150, loss: 0.012436583638191223
step: 160, loss: 0.0071983132511377335
step: 170, loss: 0.011308626271784306
step: 180, loss: 0.058091480284929276
step: 190, loss: 0.03626076877117157
step: 200, loss: 0.05527055263519287
step: 210, loss: 0.009258043952286243
step: 220, loss: 0.030655989423394203
step: 230, loss: 0.01619948446750641
step: 240, loss: 0.011154994368553162
step: 250, loss: 0.23228521645069122
step: 260, loss: 0.014348427765071392
step: 270, loss: 0.03249972686171532
step: 280, loss: 0.0033625659998506308
step: 290, loss: 0.06198284402489662
step: 300, loss: 0.0015589360846206546
step: 310, loss: 0.00438386342674494
step: 320, loss: 0.10445358604192734
step: 330, loss: 0.030514324083924294
step: 340, loss: 0.006163613870739937
step: 350, loss: 0.1478196680545807
step: 360, loss: 0.012770509347319603
step: 370, loss: 0.003074455773457885
step: 380, loss: 0.08624788373708725
step: 390, loss: 0.00045526662142947316
step: 400, loss: 0.009578254073858261
step: 410, loss: 0.010417494922876358
step: 420, loss: 0.01684717647731304
step: 430, loss: 0.036095306277275085
```

```
step: 440, loss: 0.0024440851993858814
step: 450, loss: 0.03743412718176842
step: 460, loss: 0.05198727175593376
step: 470, loss: 0.0013672946952283382
step: 480, loss: 0.004504505079239607
step: 490, loss: 0.07355892658233643
step: 500, loss: 0.03632328659296036
step: 510, loss: 0.09450748562812805
step: 520, loss: 0.022750386968255043
step: 530, loss: 0.16245193779468536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9435897435897437, f1=0.9380530973451328,
best_f1=0.9380530973451328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002667633118107915
step: 10, loss: 0.058353520929813385
step: 20, loss: 0.006775300949811935
step: 30, loss: 0.04068616405129433
step: 40, loss: 0.006591958925127983
step: 50, loss: 0.0995962917804718
step: 60, loss: 0.09080573916435242
step: 70, loss: 0.002714497037231922
step: 80, loss: 0.04688965156674385
step: 90, loss: 0.1029822826385498
step: 100, loss: 0.10474710911512375
step: 110, loss: 0.021972214803099632
step: 120, loss: 0.19122563302516937
step: 130, loss: 0.03938116505742073
step: 140, loss: 0.10309404134750366
step: 150, loss: 0.04898275062441826
step: 160, loss: 0.034449514001607895
step: 170, loss: 0.1684150993824005
step: 180, loss: 0.005361373536288738
step: 190, loss: 0.004743825178593397
step: 200, loss: 0.0125714261084795
step: 210, loss: 0.03327922150492668
step: 220, loss: 0.003293299814686179
step: 230, loss: 0.008896972984075546
step: 240, loss: 0.014566185884177685
step: 250, loss: 0.37867650389671326
```

```
step: 260, loss: 0.08692009001970291
step: 270, loss: 0.03189439699053764
step: 280, loss: 0.006509196013212204
step: 290, loss: 0.01859954372048378
step: 300, loss: 0.026334799826145172
step: 310, loss: 0.0550619401037693
step: 320, loss: 0.026584723964333534
step: 330, loss: 0.004093850962817669
step: 340, loss: 0.018461482599377632
step: 350, loss: 0.0022186855785548687
step: 360, loss: 0.003022590419277549
step: 370, loss: 0.001320741605013609
step: 380, loss: 0.0018469695933163166
step: 390, loss: 0.013088660314679146
step: 400, loss: 0.006104452535510063
step: 410, loss: 0.13819649815559387
step: 420, loss: 0.22454240918159485
step: 430, loss: 0.02459658496081829
step: 440, loss: 0.004256246145814657
step: 450, loss: 0.1292690485715866
step: 460, loss: 0.04557688906788826
step: 470, loss: 0.019418595358729362
step: 480, loss: 0.009120460599660873
step: 490, loss: 0.015446078963577747
step: 500, loss: 0.0077949450351297855
step: 510, loss: 0.00475173257291317
step: 520, loss: 0.030735256150364876
step: 530, loss: 0.04448439180850983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9481481481481482, f1=0.940959409594096,
best_f1=0.940959409594096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07002425938844681
step: 10, loss: 0.00043491602991707623
step: 20, loss: 0.0020065016578882933
step: 30, loss: 0.0019277187529951334
step: 40, loss: 0.008623739704489708
step: 50, loss: 0.0031249872408807278
step: 60, loss: 0.03301848843693733
step: 70, loss: 0.0009500084561295807
```

```
step: 80, loss: 0.025380633771419525
step: 90, loss: 0.008567322976887226
step: 100, loss: 0.004129223059862852
step: 110, loss: 0.012781553901731968
step: 120, loss: 0.24500606954097748
step: 130, loss: 0.022309251129627228
step: 140, loss: 0.001520422869361937
step: 150, loss: 0.004396605305373669
step: 160, loss: 0.026402672752738
step: 170, loss: 0.00044847093522548676
step: 180, loss: 0.04743576794862747
step: 190, loss: 0.14090821146965027
step: 200, loss: 0.005821928847581148
step: 210, loss: 0.0036538480781018734
step: 220, loss: 0.005274753086268902
step: 230, loss: 0.012029018253087997
step: 240, loss: 0.0007604667916893959
step: 250, loss: 0.16515730321407318
step: 260, loss: 0.03716890886425972
step: 270, loss: 0.009237300604581833
step: 280, loss: 0.00066615101592615247
step: 290, loss: 0.0062393867410719395
step: 300, loss: 0.03464198112487793
step: 310, loss: 0.09418626129627228
step: 320, loss: 0.0029290015809237957
step: 330, loss: 0.007005218416452408
step: 340, loss: 0.020081980153918266
step: 350, loss: 0.001018468290567398
step: 360, loss: 0.10710697621107101
step: 370, loss: 0.04894501715898514
step: 380, loss: 0.0029867403320920944
step: 390, loss: 0.003488404443487525
step: 400, loss: 0.003206158522516489
step: 410, loss: 0.00029842060757800937
step: 420, loss: 0.000766065320931375
step: 430, loss: 0.007739711087197065
step: 440, loss: 0.00032530955667912066
step: 450, loss: 0.25334233045578003
step: 460, loss: 0.0020627614576369524
step: 470, loss: 0.023723646998405457
step: 480, loss: 0.014004168100655079
step: 490, loss: 0.03440892696380615
step: 500, loss: 0.022509977221488953
step: 510, loss: 0.24602198600769043
step: 520, loss: 0.002932619536295533
step: 530, loss: 0.004546649754047394
/usr/local/lib/python3.7/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=0.944157672454247, f1=0.9390815370196813,
best_f1=0.940959409594096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015511849196627736
step: 10, loss: 0.025199508294463158
step: 20, loss: 0.0041193850338459015
step: 30, loss: 0.005026810336858034
step: 40, loss: 0.010302674956611974
step: 50, loss: 0.0026987851597368717
step: 60, loss: 0.004064635373651981
step: 70, loss: 0.0007660443079657853
step: 80, loss: 0.0025109953712671995
step: 90, loss: 0.0009787660092115402
step: 100, loss: 0.07526090741157532
step: 110, loss: 0.003284877398982644
step: 120, loss: 0.0013054553419351578
step: 130, loss: 0.00019424056517891586
step: 140, loss: 0.0011380859650671482
step: 150, loss: 0.00267621548846364
step: 160, loss: 0.0012274235486984253
step: 170, loss: 0.016278201714158058
step: 180, loss: 0.042906370013952255
step: 190, loss: 0.050351861864328384
step: 200, loss: 0.0003555816365405917
step: 210, loss: 0.00047000497579574585
step: 220, loss: 0.0003812727809417993
step: 230, loss: 0.0007653162465430796
step: 240, loss: 0.057215191423892975
step: 250, loss: 0.012011871673166752
step: 260, loss: 0.002770510269328952
step: 270, loss: 0.0025147725827991962
step: 280, loss: 0.007276608608663082
step: 290, loss: 0.03257235884666443
step: 300, loss: 0.10410937666893005
step: 310, loss: 0.001747378264553845
step: 320, loss: 0.01421814039349556
step: 330, loss: 0.0007104742689989507
step: 340, loss: 0.0020997533574700356
step: 350, loss: 0.002079112222418189
step: 360, loss: 0.040732868015766144
step: 370, loss: 0.021456332877278328
step: 380, loss: 0.004483741708099842
step: 390, loss: 0.020179476588964462
step: 400, loss: 0.01449370477348566
step: 410, loss: 0.02087625116109848
```

```
step: 420, loss: 0.008994481526315212
step: 430, loss: 0.0054157269187271595
step: 440, loss: 0.004061146639287472
step: 450, loss: 0.0012891514925286174
step: 460, loss: 0.0006678941426798701
step: 470, loss: 0.09631519019603729
step: 480, loss: 0.006792785134166479
step: 490, loss: 0.014070202596485615
step: 500, loss: 0.0012910317163914442
step: 510, loss: 0.0007566760177724063
step: 520, loss: 0.0007234050426632166
step: 530, loss: 0.0036684514489024878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9489655172413793, f1=0.9464040311497939,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016124306712299585
step: 10, loss: 0.0002687669184524566
step: 20, loss: 6.423769082175568e-05
step: 30, loss: 0.0023065435234457254
step: 40, loss: 0.20386911928653717
step: 50, loss: 0.0002951752976514399
step: 60, loss: 0.031193384900689125
step: 70, loss: 0.023265965282917023
step: 80, loss: 0.008206215687096119
step: 90, loss: 0.00019677015370689332
step: 100, loss: 0.019954144954681396
step: 110, loss: 0.0005521985003724694
step: 120, loss: 0.0010018089087679982
step: 130, loss: 0.004021646920591593
step: 140, loss: 0.0006016194238327444
step: 150, loss: 0.015207215212285519
step: 160, loss: 0.0009825596353039145
step: 170, loss: 0.17135483026504517
step: 180, loss: 0.0034742336720228195
step: 190, loss: 0.02223467268049717
step: 200, loss: 0.003776188939809799
step: 210, loss: 0.04353015497326851
step: 220, loss: 0.004474145360291004
step: 230, loss: 0.013483253307640553
```

```
step: 240, loss: 0.03259310871362686
step: 250, loss: 0.003585283411666751
step: 260, loss: 0.010170229710638523
step: 270, loss: 0.035911574959754944
step: 280, loss: 0.002546170959249139
step: 290, loss: 0.0016322721494361758
step: 300, loss: 0.00019836881256196648
step: 310, loss: 0.0004759605508297682
step: 320, loss: 0.0017491745529696345
step: 330, loss: 0.00013041780039202422
step: 340, loss: 0.011048384942114353
step: 350, loss: 0.00036402710247784853
step: 360, loss: 0.005938112270087004
step: 370, loss: 0.11951020359992981
step: 380, loss: 0.0024473306257277727
step: 390, loss: 0.06728876382112503
step: 400, loss: 0.022805865854024887
step: 410, loss: 0.005661592818796635
step: 420, loss: 0.0006863278686068952
step: 430, loss: 0.026147276163101196
step: 440, loss: 0.17980429530143738
step: 450, loss: 0.005465014837682247
step: 460, loss: 0.0028141967486590147
step: 470, loss: 0.1667764037847519
step: 480, loss: 0.035786811262369156
step: 490, loss: 0.0022264982108026743
step: 500, loss: 0.0024022054858505726
step: 510, loss: 0.014755829237401485
step: 520, loss: 0.0020997575484216213
step: 530, loss: 0.0002848786534741521
/usr/local/lib/python3.7/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=0.9458955223880597, f1=0.9437470943747095,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010163454571173884
step: 10, loss: 0.04399620741605759
step: 20, loss: 0.00034715404035523534
step: 30, loss: 0.07498244196176529
step: 40, loss: 0.03286600485444069
step: 50, loss: 0.0008178808493539691
step: 60, loss: 0.007166923955082893
step: 70, loss: 0.002339487662538886
step: 80, loss: 0.009590649046003819
step: 90, loss: 0.022049197927117348
```

```
step: 100, loss: 0.0004193554923404008
step: 110, loss: 0.011178073473274708
step: 120, loss: 0.010835018008947372
step: 130, loss: 0.00041706516640260816
step: 140, loss: 0.005926144775003195
step: 150, loss: 0.001957886153832078
step: 160, loss: 0.024728205054998398
step: 170, loss: 0.006738840602338314
step: 180, loss: 0.00386408856138587
step: 190, loss: 0.022980468347668648
step: 200, loss: 0.0017292582197114825
step: 210, loss: 0.009893575683236122
step: 220, loss: 0.0003279367520008236
step: 230, loss: 0.0003402080910746008
step: 240, loss: 0.01984076388180256
step: 250, loss: 0.025264205411076546
step: 260, loss: 0.00015932106180116534
step: 270, loss: 0.06780566275119781
step: 280, loss: 0.0035792782437056303
step: 290, loss: 0.000537946296390146
step: 300, loss: 0.0001706054317764938
step: 310, loss: 0.0643237829208374
step: 320, loss: 0.00024061418662313372
step: 330, loss: 0.0033693606965243816
step: 340, loss: 0.00842282548546791
step: 350, loss: 0.02934722788631916
step: 360, loss: 0.008702443912625313
step: 370, loss: 0.017413070425391197
step: 380, loss: 0.0026927071157842875
step: 390, loss: 0.0017410224536433816
step: 400, loss: 0.018879011273384094
step: 410, loss: 0.03544454276561737
step: 420, loss: 0.00034846417838707566
step: 430, loss: 0.0889444500207901
step: 440, loss: 0.043567392975091934
step: 450, loss: 0.0160736832767725
step: 460, loss: 8.67267808644101e-05
step: 470, loss: 0.00018586312944535166
step: 480, loss: 0.0010849767131730914
step: 490, loss: 0.004507578443735838
step: 500, loss: 0.0030006791930645704
step: 510, loss: 0.00012028328637825325
step: 520, loss: 0.0018150742398574948
step: 530, loss: 0.06925405561923981
/usr/local/lib/python3.7/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=0.9477577438742487, f1=0.9430291801760075,
```

```
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031744036823511124
step: 10, loss: 0.00023021070228423923
step: 20, loss: 0.0003257423231843859
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.001355347572825849
step: 40, loss: 0.0004220232949592173
step: 50, loss: 0.0002855985949281603
step: 60, loss: 0.00128093920648098
step: 70, loss: 0.00016309910279233009
step: 80, loss: 0.0010857128072530031
step: 90, loss: 0.008589081466197968
step: 100, loss: 0.0053495559841394424
step: 110, loss: 0.018217355012893677
step: 120, loss: 0.001617880305275321
step: 130, loss: 0.003808560548350215
step: 140, loss: 0.0067015946842730045
step: 150, loss: 0.00039561514859087765
step: 160, loss: 0.030935704708099365
step: 170, loss: 0.0025316474493592978
step: 180, loss: 0.006226326804608107
step: 190, loss: 7.61063420213759e-05
step: 200, loss: 0.00016348998178727925
step: 210, loss: 0.0008989329217001796
step: 220, loss: 0.000401239754864946
step: 230, loss: 0.0009116222499869764
step: 240, loss: 0.01231573149561882
step: 250, loss: 0.10221390426158905
step: 260, loss: 0.018474895507097244
step: 270, loss: 0.0008340837084688246
step: 280, loss: 0.0012864447198808193
step: 290, loss: 0.010435031726956367
step: 300, loss: 0.005048342049121857
step: 310, loss: 0.048130832612514496
step: 320, loss: 0.004312374163419008
step: 330, loss: 0.000784913485404104
step: 340, loss: 0.00014644522161688656
step: 350, loss: 0.024738546460866928
step: 360, loss: 0.0005892417975701392
step: 370, loss: 0.0003574688162188977
step: 380, loss: 0.0017467131838202477
step: 390, loss: 0.0009377344977110624
step: 400, loss: 0.001702392939478159
step: 410, loss: 0.00645663496106863
step: 420, loss: 0.005177557468414307
```

```
step: 430, loss: 0.0003781129780691117
step: 440, loss: 0.0001181369589176029
step: 450, loss: 0.002969199325889349
step: 460, loss: 0.0026179824490100145
step: 470, loss: 0.02284880168735981
step: 480, loss: 0.002290224190801382
step: 490, loss: 0.05745347589254379
step: 500, loss: 0.11528930813074112
step: 510, loss: 0.000812262762337923
step: 520, loss: 0.012686326168477535
step: 530, loss: 0.0016510082641616464
/usr/local/lib/python3.7/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=0.9441624365482235, f1=0.9413394919168592,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003042574389837682
step: 10, loss: 0.05311593785881996
step: 20, loss: 0.047346003353595734
step: 30, loss: 0.0016991252778097987
step: 40, loss: 0.004942391533404589
step: 50, loss: 0.0003803749568760395
step: 60, loss: 0.0007976973429322243
step: 70, loss: 0.00024171595578081906
step: 80, loss: 0.0032682535238564014
step: 90, loss: 0.0031146002002060413
step: 100, loss: 0.0023983800783753395
step: 110, loss: 0.00011874296615133062
step: 120, loss: 0.00034924791543744504
step: 130, loss: 0.007861833088099957
step: 140, loss: 3.826079046120867e-05
step: 150, loss: 0.0001704133755993098
step: 160, loss: 0.0015360096003860235
step: 170, loss: 0.011842411942780018
step: 180, loss: 0.00012638197222258896
step: 190, loss: 0.00021609691611956805
step: 200, loss: 0.0012637614272534847
step: 210, loss: 0.00171557255089283
step: 220, loss: 0.003530775895342231
step: 230, loss: 0.005510579328984022
step: 240, loss: 0.010129803791642189
step: 250, loss: 0.00036595918936654925
step: 260, loss: 0.0010991625022143126
step: 270, loss: 0.0007905091624706984
step: 280, loss: 0.002429565414786339
```

```
step: 290, loss: 0.0006753629422746599
step: 300, loss: 3.519942765706219e-05
step: 310, loss: 0.0032018464989960194
step: 320, loss: 0.11946425586938858
step: 330, loss: 6.262827082537115e-05
step: 340, loss: 0.004392441362142563
step: 350, loss: 0.020671753212809563
step: 360, loss: 0.025224409997463226
step: 370, loss: 0.0015459290007129312
step: 380, loss: 0.0032759150490164757
step: 390, loss: 0.0013373091351240873
step: 400, loss: 0.0001744577893987298
step: 410, loss: 0.005702023860067129
step: 420, loss: 0.0025550154969096184
step: 430, loss: 0.003033897839486599
step: 440, loss: 0.0013208403252065182
step: 450, loss: 0.0012031419901177287
step: 460, loss: 0.0004720432043541223
step: 470, loss: 0.00047346361679956317
step: 480, loss: 0.001006751786917448
step: 490, loss: 9.994069114327431e-05
step: 500, loss: 0.0007808583322912455
step: 510, loss: 0.0004459662886802107
step: 520, loss: 0.0034656766802072525
step: 530, loss: 0.00193046953063453582
/usr/local/lib/python3.7/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=0.9470260223048327, f1=0.9455053563111318,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029082412365823984
step: 10, loss: 0.0006118071032688022
step: 20, loss: 0.005720025859773159
step: 30, loss: 4.7540605009999126e-05
step: 40, loss: 0.015959039330482483
step: 50, loss: 0.00011738170724129304
step: 60, loss: 2.356935692660045e-05
step: 70, loss: 0.0002954851952381432
step: 80, loss: 0.002355766948312521
step: 90, loss: 0.00015709757281001657
step: 100, loss: 0.2837276756763458
step: 110, loss: 0.00013499513443093747
step: 120, loss: 0.000680529628880322
step: 130, loss: 8.707301458343863e-05
step: 140, loss: 0.07235214114189148
```

```
step: 150, loss: 0.0022886772640049458
step: 160, loss: 0.024771476164460182
step: 170, loss: 7.408763485727832e-05
step: 180, loss: 0.0006982598570175469
step: 190, loss: 0.00015253169112838805
step: 200, loss: 0.0016369092045351863
step: 210, loss: 0.0025465337093919516
step: 220, loss: 5.099003465147689e-05
step: 230, loss: 0.0007006392697803676
step: 240, loss: 7.383854244835675e-05
step: 250, loss: 2.808391764119733e-05
step: 260, loss: 9.401199349667877e-05
step: 270, loss: 0.011895889416337013
step: 280, loss: 0.002261542482301593
step: 290, loss: 0.0004716241965070367
step: 300, loss: 0.0003764393331948668
step: 310, loss: 0.006900846026837826
step: 320, loss: 0.01853710040450096
step: 330, loss: 0.002488830592483282
step: 340, loss: 0.0006527221412397921
step: 350, loss: 3.8539525121450424e-05
step: 360, loss: 0.013364985585212708
step: 370, loss: 0.0052838195115327835
step: 380, loss: 0.0002742684737313539
step: 390, loss: 0.030860217288136482
step: 400, loss: 3.0192662961781025e-05
step: 410, loss: 0.00019496853929013014
step: 420, loss: 0.0009062481112778187
step: 430, loss: 0.0015494307735934854
step: 440, loss: 0.01173496339470148
step: 450, loss: 0.00027196307200938463
step: 460, loss: 0.0009464171016588807
step: 470, loss: 0.0016960377106443048
step: 480, loss: 0.008073526434600353
step: 490, loss: 0.019726062193512917
step: 500, loss: 0.0010059773921966553
step: 510, loss: 0.011282532475888729
step: 520, loss: 0.007292834110558033
step: 530, loss: 0.0017796573229134083
/usr/local/lib/python3.7/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=0.9484820607175714, f1=0.9430147058823529,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008181845187209547
```

```
step: 10, loss: 0.0004816631262656301
step: 20, loss: 0.0005640804884023964
step: 30, loss: 4.3745509174186736e-05
step: 40, loss: 0.0009318552911281586
step: 50, loss: 0.004497295245528221
step: 60, loss: 0.0008283042116090655
step: 70, loss: 0.08721847832202911
step: 80, loss: 0.004280542954802513
step: 90, loss: 0.000254279060754925
step: 100, loss: 0.00017443856631871313
step: 110, loss: 0.000374165567336604
step: 120, loss: 0.0007025680970400572
step: 130, loss: 0.0019281633431091905
step: 140, loss: 0.00022638554219156504
step: 150, loss: 0.002997568342834711
step: 160, loss: 0.00016619751113466918
step: 170, loss: 0.08680545538663864
step: 180, loss: 0.0004254742816556245
step: 190, loss: 8.53022065712139e-05
step: 200, loss: 0.000722218886949122
step: 210, loss: 2.1263587768771686e-05
step: 220, loss: 0.004277332220226526
step: 230, loss: 0.003295426955446601
step: 240, loss: 0.0007175438804551959
step: 250, loss: 0.0015219154302030802
step: 260, loss: 1.933757812366821e-05
step: 270, loss: 1.9281813365523703e-05
step: 280, loss: 0.013683309778571129
step: 290, loss: 2.2916712623555213e-05
step: 300, loss: 0.0008257530280388892
step: 310, loss: 0.00062646868173033
step: 320, loss: 8.688671368872747e-05
step: 330, loss: 0.0014012180035933852
step: 340, loss: 0.0077467309311032295
step: 350, loss: 0.00045493600191548467
step: 360, loss: 0.006769995670765638
step: 370, loss: 0.006115152966231108
step: 380, loss: 0.00021605541405733675
step: 390, loss: 8.245973731391132e-05
step: 400, loss: 0.00015392208297271281
step: 410, loss: 2.144577410945203e-05
step: 420, loss: 0.032901886850595474
step: 430, loss: 0.0023894135374575853
step: 440, loss: 4.579612141242251e-05
step: 450, loss: 0.00012023356975987554
step: 460, loss: 0.007489049341529608
step: 470, loss: 0.0006309088203124702
step: 480, loss: 2.2257649106904864e-05
```

```
step: 490, loss: 0.003311803797259927
step: 500, loss: 0.0005404372932389379
step: 510, loss: 0.00046388880582526326
step: 520, loss: 0.0010662435088306665
step: 530, loss: 0.0003821802674792707
/usr/local/lib/python3.7/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=0.9465364946536494, f1=0.9437470943747095,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016515729657839984
step: 10, loss: 0.0020499243400990963
step: 20, loss: 3.56349628418684e-05
step: 30, loss: 9.918786963680759e-05
step: 40, loss: 0.00044653742224909365
step: 50, loss: 0.016515163704752922
step: 60, loss: 0.0007236535893753171
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.34785303473472595
step: 80, loss: 0.0037269622553139925
step: 90, loss: 3.1166451663011685e-05
step: 100, loss: 0.005705798976123333
step: 110, loss: 0.0034588503185659647
step: 120, loss: 0.004762730095535517
step: 130, loss: 0.0005328061524778605
step: 140, loss: 0.0045684101060032845
step: 150, loss: 0.002373317489400506
step: 160, loss: 0.0010811180109158158
step: 170, loss: 0.00040707492735236883
step: 180, loss: 0.12410424649715424
step: 190, loss: 0.03319188952445984
step: 200, loss: 0.0013317539123818278
step: 210, loss: 0.020427796989679337
step: 220, loss: 0.00028804628527723253
step: 230, loss: 0.0002206944045610726
step: 240, loss: 0.0008801709627732635
step: 250, loss: 0.003266852581873536
step: 260, loss: 0.0016893307911232114
step: 270, loss: 0.006509415805339813
step: 280, loss: 0.0006741018733009696
step: 290, loss: 0.000795061350800097
step: 300, loss: 0.00029246206395328045
step: 310, loss: 0.12020336091518402
step: 320, loss: 0.00013158466026652604
step: 330, loss: 0.0005546393222175539
```

```
step: 340, loss: 0.00028919894248247147
step: 350, loss: 0.00013020697224419564
step: 360, loss: 0.00025173721951432526
step: 370, loss: 0.0016003426862880588
step: 380, loss: 0.007649743929505348
step: 390, loss: 0.022088371217250824
step: 400, loss: 0.0038297639694064856
step: 410, loss: 2.559560743975453e-05
step: 420, loss: 3.5961442335974425e-05
step: 430, loss: 0.012084534391760826
step: 440, loss: 0.00045330068678595126
step: 450, loss: 0.00010237583046546206
step: 460, loss: 0.022068334743380547
step: 470, loss: 0.0005838746437802911
step: 480, loss: 0.001158901839517057
step: 490, loss: 4.461020580492914e-05
step: 500, loss: 0.02111993171274662
step: 510, loss: 0.0011060629040002823
step: 520, loss: 0.0002282060740981251
step: 530, loss: 0.0006861412548460066
/usr/local/lib/python3.7/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=0.9465861588481189, f1=0.9460465116279071,
best_f1=0.9464040311497939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.4328555809915997e-05
step: 10, loss: 0.00023230862279888242
step: 20, loss: 0.0013756509870290756
step: 30, loss: 0.010727571323513985
step: 40, loss: 0.0011511736083775759
step: 50, loss: 0.0023540505208075047
step: 60, loss: 0.00036994231049902737
step: 70, loss: 0.0010480147320777178
step: 80, loss: 2.525936179154087e-05
step: 90, loss: 0.0002037588128587231
step: 100, loss: 2.769432103377767e-05
step: 110, loss: 0.0005291744600981474
step: 120, loss: 0.00031244748970493674
step: 130, loss: 0.0009287996217608452
step: 140, loss: 0.00047022791113704443
step: 150, loss: 7.900955097284168e-05
step: 160, loss: 0.00018309641745872796
step: 170, loss: 0.00042955964454449713
step: 180, loss: 0.014516347087919712
step: 190, loss: 0.0010907112155109644
```

```
step: 200, loss: 0.0007024743244983256
step: 210, loss: 0.00097547413315624
step: 220, loss: 1.4651353922090493e-05
step: 230, loss: 0.0009758305968716741
step: 240, loss: 5.214768316363916e-05
step: 250, loss: 0.004330188501626253
step: 260, loss: 1.027428970701294e-05
step: 270, loss: 0.00021611845295410603
step: 280, loss: 9.198048792313784e-05
step: 290, loss: 0.00010575100895948708
step: 300, loss: 0.0009462334564886987
step: 310, loss: 3.005312646564562e-05
step: 320, loss: 0.00011998658737866208
step: 330, loss: 0.0004272715887054801
step: 340, loss: 3.803958679782227e-05
step: 350, loss: 0.0011286073131486773
step: 360, loss: 0.0007028306135907769
step: 370, loss: 0.04463263601064682
step: 380, loss: 0.0005578668788075447
step: 390, loss: 0.00026298430748283863
step: 400, loss: 0.011465256102383137
step: 410, loss: 0.00021063766325823963
step: 420, loss: 0.00025324197486042976
step: 430, loss: 1.6115309335873462e-05
step: 440, loss: 0.004166353028267622
step: 450, loss: 0.0009569150279276073
step: 460, loss: 0.008377193473279476
step: 470, loss: 1.738538412610069e-05
step: 480, loss: 0.0004484129894990474
step: 490, loss: 0.0018105709459632635
step: 500, loss: 0.0008749391417950392
step: 510, loss: 3.278621443314478e-05
step: 520, loss: 1.678568332863506e-05
step: 530, loss: 0.0001108460346586071
/usr/local/lib/python3.7/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=0.9453197405004634, f1=0.9465364946536494,
best_f1=0.9464040311497939
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 [00:32, 174.08it/s]
load_f1 = 0.9507139567019807
real_f1 = 0.9508497932935233
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:29, 148.29it/s]

2   BASELINE

2.1   BASELINE STRUCTURED

2.1.1   Beer - Running the matcher

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

```
    --finetuning \
    --lm roberta \
    --fp16 \
    --save_model

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

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

Selected optimization level 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: 0.42390382289886475
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.3384615384615385, f1=0.2898550724637682,
best_f1=0.2898550724637682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3873877227306366
/usr/local/lib/python3.7/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=0.32558139534883723, f1=0.30952380952380953,
best_f1=0.2898550724637682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4205045700073242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.46511627906976755, f1=0.4705882352941177,
best_f1=0.4705882352941177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21751780807971954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

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=0.5714285714285714, 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: 0.22293953597545624
/usr/local/lib/python3.7/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=0.4761904761904762, f1=0.5, 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: 0.22154931724071503
/usr/local/lib/python3.7/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=0.5217391304347825, f1=0.4680851063829786,
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: 0.39110514521598816
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.56, f1=0.6428571428571429, 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: 0.23623912036418915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.6250000000000001,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11533186584711075

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7096774193548386, f1=0.5517241379310344,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2389116883277893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5714285714285714,
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: 0.07713271677494049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.6428571428571429,
best_f1=0.6428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028094511479139328
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6451612903225806,
best_f1=0.6428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03831807151436806
```

```
/usr/local/lib/python3.7/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=0.6206896551724138,
best_f1=0.6428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011318630538880825
/usr/local/lib/python3.7/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=0.6206896551724138,
best_f1=0.6428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015275615267455578
/usr/local/lib/python3.7/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=0.6206896551724138,
best_f1=0.6428571428571429
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, 104972.95it/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 = 0.8
real_f1 = 0.7407407407407408
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 205.42it/s]
```

2.1.2 DBLP-ACM - Running the matcher

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

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 282kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 6.30MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 3.66MB/s]
```

```
Downloading: 100% 501M/501M [00:10<00:00, 45.6MB/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: 0.5407029986381531
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.458955854177475
step: 20, loss: 0.4525727927684784
step: 30, loss: 0.31337687373161316
step: 40, loss: 0.38228970766067505
step: 50, loss: 0.5834799408912659
step: 60, loss: 0.4627467393875122
step: 70, loss: 0.4181715250015259
step: 80, loss: 0.5170038342475891
step: 90, loss: 0.2528969645500183
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.19546136260032654
step: 110, loss: 0.09395655244588852
step: 120, loss: 0.1969011276960373
step: 130, loss: 0.032674338668584824
step: 140, loss: 0.016945382580161095
step: 150, loss: 0.11542721092700958
step: 160, loss: 0.027295183390378952
step: 170, loss: 0.09969130158424377
step: 180, loss: 0.23370395600795746
step: 190, loss: 0.09379993379116058
step: 200, loss: 0.035644158720970154
step: 210, loss: 0.02421395108103752
step: 220, loss: 0.11591622233390808
step: 230, loss: 0.006837813649326563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.96353591160221, f1=0.9644444444444443,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004543609917163849
step: 10, loss: 0.09160922467708588
step: 20, loss: 0.024108529090881348
step: 30, loss: 0.06254881620407104
step: 40, loss: 0.010374440811574459
step: 50, loss: 0.03630821406841278
step: 60, loss: 0.012477615848183632
step: 70, loss: 0.008107136934995651
step: 80, loss: 0.011790875345468521
step: 90, loss: 0.0133594935759902
step: 100, loss: 0.06511504203081131
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 110, loss: 0.15768566727638245
step: 120, loss: 0.06518440693616867
step: 130, loss: 0.03474687039852142
step: 140, loss: 0.15432722866535187
step: 150, loss: 0.1385224163532257
step: 160, loss: 0.2898031771183014
step: 170, loss: 0.0049456385895609856
step: 180, loss: 0.01412897277623415
step: 190, loss: 0.017920857295393944
```

```
step: 200, loss: 0.040094571992754936
step: 210, loss: 0.02014089748263359
step: 220, loss: 0.0009296108037233353
step: 230, loss: 0.0028925342485308647
/usr/local/lib/python3.7/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=0.9300699300699301, f1=0.9445727482678984,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19837287068367004
step: 10, loss: 0.05824713781476021
step: 20, loss: 0.062370460480451584
step: 30, loss: 0.02284439280629158
step: 40, loss: 0.029673676937818527
step: 50, loss: 0.017136225476861
step: 60, loss: 0.003729697549715638
step: 70, loss: 0.0018496045377105474
step: 80, loss: 0.03562057018280029
step: 90, loss: 0.015031270682811737
step: 100, loss: 0.004266114439815283
step: 110, loss: 0.0180410947650671
step: 120, loss: 0.0009822737192735076
step: 130, loss: 0.006563854869455099
step: 140, loss: 0.0012426907196640968
step: 150, loss: 0.05231266841292381
step: 160, loss: 0.0025013431441038847
step: 170, loss: 0.004028154071420431
step: 180, loss: 0.01924981363117695
step: 190, loss: 0.052410777658224106
step: 200, loss: 0.01707615703344345
step: 210, loss: 0.001855584792792797
step: 220, loss: 0.011488660238683224
step: 230, loss: 0.02861616015434265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9707865168539327, f1=0.9730337078651685,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.02039012312889099
step: 10, loss: 0.0012204116210341454
step: 20, loss: 0.013457468710839748
step: 30, loss: 0.056224703788757324
step: 40, loss: 0.17543277144432068
step: 50, loss: 0.07815714180469513
step: 60, loss: 0.0025490059051662683
step: 70, loss: 0.012122239917516708
step: 80, loss: 0.0008118956466205418
step: 90, loss: 0.08581985533237457
step: 100, loss: 0.0031506808008998632
step: 110, loss: 0.01709963008761406
step: 120, loss: 0.029943037778139114
step: 130, loss: 0.0390593558549881
step: 140, loss: 0.00055063463514670073
step: 150, loss: 0.00044130015885457397
step: 160, loss: 0.0018902334850281477
step: 170, loss: 0.09021832793951035
step: 180, loss: 0.07286254316568375
step: 190, loss: 0.023940762504935265
step: 200, loss: 0.007344034966081381
step: 210, loss: 0.011268545873463154
step: 220, loss: 0.0004111565067432821
step: 230, loss: 0.007536699995398521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9821029082774049, 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: 0.0022770031355321407
step: 10, loss: 0.0017412889283150434
step: 20, loss: 0.005958231166005135
step: 30, loss: 0.0008594018290750682
step: 40, loss: 0.024763798341155052
step: 50, loss: 0.005586158484220505
step: 60, loss: 0.0008860945818014443
step: 70, loss: 0.001654652995057404
step: 80, loss: 0.010608638636767864
step: 90, loss: 0.0645434707403183
step: 100, loss: 0.0005020466633141041
step: 110, loss: 0.0007174519123509526
```

```
step: 120, loss: 0.010149175301194191
step: 130, loss: 0.00023784697987139225
step: 140, loss: 0.08828141540288925
step: 150, loss: 0.15069670975208282
step: 160, loss: 0.00029116077348589897
step: 170, loss: 0.03402663767337799
step: 180, loss: 0.003204015316441655
step: 190, loss: 0.004278122913092375
step: 200, loss: 0.014221093617379665
step: 210, loss: 0.006203420460224152
step: 220, loss: 0.003529689507558942
step: 230, loss: 0.08906658738851547
/usr/local/lib/python3.7/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=0.9742441209406495, f1=0.9787234042553192,
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: 0.001000504707917571
step: 10, loss: 0.003325662575662136
step: 20, loss: 0.002853184472769499
step: 30, loss: 0.0022822644095867872
step: 40, loss: 0.0004946927074342966
step: 50, loss: 0.0005072814528830349
step: 60, loss: 0.0025454917922616005
step: 70, loss: 0.00029176496900618076
step: 80, loss: 0.000994822010397911
step: 90, loss: 0.06439710408449173
step: 100, loss: 0.00615145917981863
step: 110, loss: 0.08566148579120636
step: 120, loss: 0.00023600964050274342
step: 130, loss: 0.0002529555349610746
step: 140, loss: 0.0001705063768895343
step: 150, loss: 0.0008465864812023938
step: 160, loss: 0.02135912887752056
step: 170, loss: 0.009376230649650097
step: 180, loss: 0.0010253319051116705
step: 190, loss: 0.00027846349985333875
step: 200, loss: 0.001822657068260014
step: 210, loss: 0.000729091523680836
step: 220, loss: 0.010839096270501614
step: 230, loss: 0.0007859174511395395
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.9818594104308391,
```

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: 0.0027149890083819628
step: 10, loss: 0.0002114351955242455
step: 20, loss: 0.24429138004779816
step: 30, loss: 0.0026828530244529247
step: 40, loss: 0.004253532737493515
step: 50, loss: 0.0010193749330937862
step: 60, loss: 0.00042601730092428625
step: 70, loss: 0.0006778023671358824
step: 80, loss: 0.007972278632223606
step: 90, loss: 0.00031826275517232716
step: 100, loss: 0.0001809128007153049
step: 110, loss: 0.000141331140184775
step: 120, loss: 0.002940445439890027
step: 130, loss: 0.0029416850302368402
step: 140, loss: 0.00046236696653068066
step: 150, loss: 0.01635829359292984
step: 160, loss: 0.000499710557051003
step: 170, loss: 0.021163256838917732
step: 180, loss: 0.002372109331190586
step: 190, loss: 0.0038525559939444065
step: 200, loss: 0.0019861862529069185
step: 210, loss: 0.0021078886929899454
step: 220, loss: 0.000191581406397745
step: 230, loss: 0.0005502347485162318
/usr/local/lib/python3.7/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=0.9783845278725825, f1=0.9796380090497738,
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: 0.001498029101639986
step: 10, loss: 0.001972963334992528
step: 20, loss: 0.0007694544037804008
step: 30, loss: 0.00032535273930989206
step: 40, loss: 0.0005299923941493034
step: 50, loss: 0.00028097842005081475
step: 60, loss: 0.0002996862167492509
step: 70, loss: 0.00010892453428823501
step: 80, loss: 0.0024356315843760967
step: 90, loss: 0.00022993395396042615
step: 100, loss: 0.00018049874051939696
step: 110, loss: 0.0028497364837676287

```
step: 120, loss: 0.018394572660326958
step: 130, loss: 0.0006129380199126899
step: 140, loss: 0.0002085860469378531
step: 150, loss: 0.016018835827708244
step: 160, loss: 0.016252148896455765
step: 170, loss: 0.04582539200782776
step: 180, loss: 0.00022368523059412837
step: 190, loss: 0.009734834544360638
step: 200, loss: 0.0395040288567543
step: 210, loss: 0.0015315820928663015
step: 220, loss: 0.04269372299313545
step: 230, loss: 0.00024023093283176422
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830124575311437, f1=0.9864864864864865,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0039876424707472324
step: 10, loss: 0.004395483061671257
step: 20, loss: 0.0028364716563373804
step: 30, loss: 0.01951294019818306
step: 40, loss: 0.028801795095205307
step: 50, loss: 0.00042132657836191356
step: 60, loss: 0.042528726160526276
step: 70, loss: 0.06977953761816025
step: 80, loss: 0.00017913820920512080
step: 90, loss: 0.07573747634887695
step: 100, loss: 0.0013778938446193933
step: 110, loss: 0.010020716115832329
step: 120, loss: 5.640533345285803e-05
step: 130, loss: 0.02789396606385708
step: 140, loss: 0.0013172312173992395
step: 150, loss: 0.0030708485282957554
step: 160, loss: 0.0006062643951736391
step: 170, loss: 0.0005051498883403838
step: 180, loss: 8.379476639674976e-05
step: 190, loss: 7.885281956987455e-05
step: 200, loss: 0.0010910475393757224
step: 210, loss: 0.032884836196899414
step: 220, loss: 0.08970963209867477
step: 230, loss: 0.04841277748346329
```

```
/usr/local/lib/python3.7/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=0.9819413092550789, f1=0.9865470852017937,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028066866798326373
step: 10, loss: 0.0025903075002133846
step: 20, loss: 0.0004708845808636397
step: 30, loss: 6.382892752299085e-05
step: 40, loss: 0.00012402443098835647
step: 50, loss: 0.003503114217892289
step: 60, loss: 4.980620360583998e-05
step: 70, loss: 0.04956398904323578
step: 80, loss: 0.0005721034249290824
step: 90, loss: 4.1648399928817526e-05
step: 100, loss: 4.4477161281974986e-05
step: 110, loss: 0.0012160456972196698
step: 120, loss: 3.8569247408304363e-05
step: 130, loss: 0.0008082055719569325
step: 140, loss: 0.0006085516652092338
step: 150, loss: 0.2721739113330841
step: 160, loss: 7.893407018855214e-05
step: 170, loss: 0.004701972007751465
step: 180, loss: 0.0006009842618368566
step: 190, loss: 0.0002481224073562771
step: 200, loss: 0.0012268901336938143
step: 210, loss: 0.02164503186941147
step: 220, loss: 0.03465857729315758
step: 230, loss: 0.00014579229173250496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842342342342343, f1=0.9787709497206705,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.804542449070141e-05
step: 10, loss: 0.00015379830438178033
step: 20, loss: 0.00027528186910785735
step: 30, loss: 0.0027787480503320694
```

```
step: 40, loss: 4.029670526506379e-05
step: 50, loss: 0.00011784005619119853
step: 60, loss: 0.003615673165768385
step: 70, loss: 0.003225587075576186
step: 80, loss: 0.00045709905680269003
step: 90, loss: 0.001538593671284616
step: 100, loss: 0.0001133488112827763
step: 110, loss: 0.00010181450488744304
step: 120, loss: 8.809083374217153e-05
step: 130, loss: 8.379179780604318e-05
step: 140, loss: 0.0005327510298229754
step: 150, loss: 0.00010537832713453099
step: 160, loss: 0.0073043182492256165
step: 170, loss: 0.004253057762980461
step: 180, loss: 6.971148832235485e-05
step: 190, loss: 5.386267730500549e-05
step: 200, loss: 0.0016757047269493341
step: 210, loss: 5.5281645472859964e-05
step: 220, loss: 5.553508526645601e-05
step: 230, loss: 5.971447171759792e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9853438556933484, f1=0.9854423292273236,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.837673860602081e-05
step: 10, loss: 5.391605009208433e-05
step: 20, loss: 0.003087047953158617
step: 30, loss: 0.07345561683177948
step: 40, loss: 0.0001650671474635601
step: 50, loss: 0.0015203319489955902
step: 60, loss: 0.0016865666257217526
step: 70, loss: 0.0013623000122606754
step: 80, loss: 8.190095832105726e-05
step: 90, loss: 0.00042624236084520817
step: 100, loss: 7.606240251334384e-05
step: 110, loss: 0.00010470656707184389
step: 120, loss: 0.0005139447748661041
step: 130, loss: 0.0003114177379757166
step: 140, loss: 5.8277324569644406e-05
step: 150, loss: 6.438101991079748e-05
```

```
step: 160, loss: 4.017837636638433e-05
step: 170, loss: 0.00015410399646498263
step: 180, loss: 0.0008205900667235255
step: 190, loss: 0.008859439752995968
step: 200, loss: 0.00030939315911382437
step: 210, loss: 6.089866656111553e-05
step: 220, loss: 0.004344083368778229
step: 230, loss: 0.00010468747495906428
/usr/local/lib/python3.7/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=0.9799107142857142,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.709412339958362e-05
step: 10, loss: 3.9980950532481074e-05
step: 20, loss: 6.549481622641906e-05
step: 30, loss: 3.1696181395091116e-05
step: 40, loss: 0.00015674348105676472
step: 50, loss: 0.00010710070637287572
step: 60, loss: 7.969316357048228e-05
step: 70, loss: 0.006175968796014786
step: 80, loss: 2.7752035748562776e-05
step: 90, loss: 0.0003826302709057927
step: 100, loss: 0.0003835057723335922
step: 110, loss: 0.0007861956837587059
step: 120, loss: 6.277641659835353e-05
step: 130, loss: 4.298362182453275e-05
step: 140, loss: 0.0009213362354785204
step: 150, loss: 0.0015125063946470618
step: 160, loss: 0.06806807965040207
step: 170, loss: 4.3832849769387394e-05
step: 180, loss: 0.010824887081980705
step: 190, loss: 4.8757607146399096e-05
step: 200, loss: 8.201022137654945e-05
step: 210, loss: 7.802280015312135e-05
step: 220, loss: 4.224967415211722e-05
step: 230, loss: 0.00033429384347982705
/usr/local/lib/python3.7/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=0.9841986455981941,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.00037172797601670027
step: 10, loss: 3.576722156140022e-05
step: 20, loss: 0.0011317235184833407
step: 30, loss: 0.00045947678154334426
step: 40, loss: 3.7678953958675265e-05
step: 50, loss: 3.288864536443725e-05
step: 60, loss: 5.757948019891046e-05
step: 70, loss: 6.255926564335823e-05
step: 80, loss: 0.0003639812639448792
step: 90, loss: 0.00012098777369828895
step: 100, loss: 9.356526425108314e-05
step: 110, loss: 6.627411494264379e-05
step: 120, loss: 2.157973540306557e-05
step: 130, loss: 0.0023346347734332085
step: 140, loss: 7.352028478635475e-05
step: 150, loss: 3.630495484685525e-05
step: 160, loss: 0.00022938140318728983
step: 170, loss: 0.0003255572519265115
step: 180, loss: 3.66510430467315e-05
step: 190, loss: 3.4616761695360765e-05
step: 200, loss: 5.106163371237926e-05
step: 210, loss: 0.000337011442752555
step: 220, loss: 0.0007620382821187377
step: 230, loss: 0.0018438927363604307
/usr/local/lib/python3.7/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=0.9843400447427293,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028231754899024963
step: 10, loss: 3.822329381364398e-05
step: 20, loss: 0.0009637723560445011
step: 30, loss: 3.8275073166005313e-05
step: 40, loss: 2.827363641699776e-05
step: 50, loss: 3.0575614800909534e-05
step: 60, loss: 0.016112761572003365
step: 70, loss: 7.208411261672154e-05
step: 80, loss: 0.000493809871841222
step: 90, loss: 3.3264936064369977e-05
step: 100, loss: 0.00010309353092452511
step: 110, loss: 0.001460961066186428
step: 120, loss: 0.02866422012448311
step: 130, loss: 0.00010073795419884846
step: 140, loss: 0.0078044794499874115
step: 150, loss: 4.3180931243114173e-05
```

```
step: 160, loss: 0.017295069992542267
step: 170, loss: 2.8139178539277054e-05
step: 180, loss: 6.206270336406305e-05
step: 190, loss: 4.2824845877476037e-05
step: 200, loss: 0.0020532028283923864
step: 210, loss: 0.014361205510795116
step: 220, loss: 5.096437962492928e-05
step: 230, loss: 4.081965744262561e-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=0.9843400447427293,
best_f1=0.9854423292273236
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, 216.34it/s]
load_f1 = 0.9852440408626559
real_f1 = 0.9852440408626559
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 201.84it/s]
```

2.1.3 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 roberta \
       --fp16 \
       --save_model


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

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

```
Downloading: 100% 481/481 [00:00<00:00, 474kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 2.85MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 1.81MB/s]
Downloading: 100% 501M/501M [00:14<00:00, 35.1MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
```

```
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6439870595932007
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5520127415657043
step: 20, loss: 0.28874707221984863
step: 30, loss: 0.3316850960254669
step: 40, loss: 0.3779626190662384
step: 50, loss: 0.5527176856994629
step: 60, loss: 0.2670207619667053
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 70, loss: 0.49898242950439453
step: 80, loss: 0.2794594466686249
step: 90, loss: 0.18805114924907684
step: 100, loss: 0.3776422142982483
step: 110, loss: 0.13978518545627594
step: 120, loss: 0.22254084050655365
step: 130, loss: 0.26228830218315125
step: 140, loss: 0.08558694273233414
step: 150, loss: 0.09645283967256546
step: 160, loss: 0.280436247587204
step: 170, loss: 0.05882886052131653
step: 180, loss: 0.1276947557926178
step: 190, loss: 0.06982436776161194
step: 200, loss: 0.05285501480102539
step: 210, loss: 0.12107114493846893
step: 220, loss: 0.09684382379055023
step: 230, loss: 0.22017191350460052
step: 240, loss: 0.047845128923654556
```

```
step: 250, loss: 0.072333341038227081
step: 260, loss: 0.1830524355173111
step: 270, loss: 0.31417447328567505
step: 280, loss: 0.07867441326379776
step: 290, loss: 0.10920055210590363
step: 300, loss: 0.05604980140924454
step: 310, loss: 0.14124464988708496
step: 320, loss: 0.025152895599603653
step: 330, loss: 0.06125317141413689
step: 340, loss: 0.4498535990715027
step: 350, loss: 0.29028430581092834
step: 360, loss: 0.044594887644052505
step: 370, loss: 0.033136311918497086
step: 380, loss: 0.06589728593826294
step: 390, loss: 0.009663141332566738
step: 400, loss: 0.03810710459947586
step: 410, loss: 0.3486086428165436
step: 420, loss: 0.01697617582976818
step: 430, loss: 0.018574358895421028
step: 440, loss: 0.021243179216980934
step: 450, loss: 0.027451371774077415
step: 460, loss: 0.009246048517525196
step: 470, loss: 0.037162862718105316
step: 480, loss: 0.09779906272888184
step: 490, loss: 0.17105187475681305
step: 500, loss: 0.04224466159939766
step: 510, loss: 0.046855051070451736
step: 520, loss: 0.04608790948987007
step: 530, loss: 0.07454598695039749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9359513791491351, f1=0.9383402874362541,
best_f1=0.9383402874362541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09771553426980972
step: 10, loss: 0.060197390615940094
step: 20, loss: 0.04921790212392807
step: 30, loss: 0.06721567362546921
step: 40, loss: 0.07584970444440842
step: 50, loss: 0.043825842440128326
step: 60, loss: 0.019058488309383392
```

```
step: 70, loss: 0.016950491815805435
step: 80, loss: 0.046452417969703674
step: 90, loss: 0.06600503623485565
step: 100, loss: 0.19167345762252808
step: 110, loss: 0.008539481088519096
step: 120, loss: 0.058190278708934784
step: 130, loss: 0.010590201243758202
step: 140, loss: 0.1870916336774826
step: 150, loss: 0.05317460000514984
step: 160, loss: 0.034445714205503464
step: 170, loss: 0.017419684678316116
step: 180, loss: 0.03625383600592613
step: 190, loss: 0.027869783341884613
step: 200, loss: 0.20629051327705383
step: 210, loss: 0.03598308563232422
step: 220, loss: 0.0031735741067677736
step: 230, loss: 0.03675134479999542
step: 240, loss: 0.09560377895832062
step: 250, loss: 0.03665297478437424
step: 260, loss: 0.019225362688302994
step: 270, loss: 0.04495418444275856
step: 280, loss: 0.04082423448562622
step: 290, loss: 0.03803228959441185
step: 300, loss: 0.05080816522240639
step: 310, loss: 0.11479339748620987
step: 320, loss: 0.02159864641726017
step: 330, loss: 0.07192947715520859
step: 340, loss: 0.014353287406265736
step: 350, loss: 0.0028047726955264807
step: 360, loss: 0.1918097883462906
step: 370, loss: 0.005438939202576876
step: 380, loss: 0.2941141724586487
step: 390, loss: 0.008527819998562336
step: 400, loss: 0.04550011828541756
step: 410, loss: 0.01634819433093071
step: 420, loss: 0.032355230301618576
step: 430, loss: 0.19871222972869873
step: 440, loss: 0.0662417858839035
step: 450, loss: 0.07783589512109756
step: 460, loss: 0.053882330656051636
step: 470, loss: 0.0140338484197855
step: 480, loss: 0.0019763638265430927
step: 490, loss: 0.005790220573544502
step: 500, loss: 0.02501860074698925
step: 510, loss: 0.06778869032859802
step: 520, loss: 0.16930927336215973
step: 530, loss: 0.015567539259791374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9522041763341067, f1=0.9429234338747099,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10525721311569214
step: 10, loss: 0.016196394339203835
step: 20, loss: 0.003624377539381385
step: 30, loss: 0.04294366389513016
step: 40, loss: 0.05321069806814194
step: 50, loss: 0.06947014480829239
step: 60, loss: 0.019961973652243614
step: 70, loss: 0.015665270388126373
step: 80, loss: 0.020158644765615463
step: 90, loss: 0.04148700460791588
step: 100, loss: 0.019342219457030296
step: 110, loss: 0.04643024504184723
step: 120, loss: 0.22295372188091278
step: 130, loss: 0.15226204693317413
step: 140, loss: 0.008589483797550201
step: 150, loss: 0.01870054006576538
step: 160, loss: 0.015555229038000107
step: 170, loss: 0.013049171306192875
step: 180, loss: 0.005524074658751488
step: 190, loss: 0.005388959776610136
step: 200, loss: 0.006302269175648689
step: 210, loss: 0.006529076490551233
step: 220, loss: 0.05947902426123619
step: 230, loss: 0.10959257930517197
step: 240, loss: 0.07789857685565948
step: 250, loss: 0.19127807021141052
step: 260, loss: 0.1011296883225441
step: 270, loss: 0.007299672346562147
step: 280, loss: 0.08034197241067886
step: 290, loss: 0.003850566688925028
step: 300, loss: 0.206427201628685
step: 310, loss: 0.020482514053583145
step: 320, loss: 0.026622414588928223
step: 330, loss: 0.047536320984363556
step: 340, loss: 0.013561113737523556
step: 350, loss: 0.06531447172164917
step: 360, loss: 0.008807577192783356
```

```
step: 370, loss: 0.0073200068436563015
step: 380, loss: 0.052877623587846756
step: 390, loss: 0.009694993495941162
step: 400, loss: 0.08510676771402359
step: 410, loss: 0.053819093853235245
step: 420, loss: 0.005293852649629116
step: 430, loss: 0.01860164850950241
step: 440, loss: 0.22365321218967438
step: 450, loss: 0.02633531205356121
step: 460, loss: 0.08283635228872299
step: 470, loss: 0.005082201678305864
step: 480, loss: 0.09194324910640717
step: 490, loss: 0.05290539562702179
step: 500, loss: 0.012425235472619534
step: 510, loss: 0.056443940848112106
step: 520, loss: 0.021457169204950333
step: 530, loss: 0.024190671741962433
/usr/local/lib/python3.7/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=0.9511278195488723, f1=0.9441052137153593,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017684031277894974
step: 10, loss: 0.010313168168067932
step: 20, loss: 0.1567077487707138
step: 30, loss: 0.17574022710323334
step: 40, loss: 0.01341771800071001
step: 50, loss: 0.0498570054769516
step: 60, loss: 0.011980298906564713
step: 70, loss: 0.025784160941839218
step: 80, loss: 0.17037071287631989
step: 90, loss: 0.07078699767589569
step: 100, loss: 0.0033765695989131927
step: 110, loss: 0.06974971294403076
step: 120, loss: 0.014158040285110474
step: 130, loss: 0.07123985141515732
step: 140, loss: 0.0022385248448699713
step: 150, loss: 0.010804212652146816
step: 160, loss: 0.009691321291029453
step: 170, loss: 0.021250221878290176
step: 180, loss: 0.06794843822717667
step: 190, loss: 0.02696036547422409
step: 200, loss: 0.016736460849642754
step: 210, loss: 0.0032852725125849247
step: 220, loss: 0.06782907247543335
```

```
step: 230, loss: 0.0037813466042280197
step: 240, loss: 0.005612076725810766
step: 250, loss: 0.06371266394853592
step: 260, loss: 0.001034098444506526
step: 270, loss: 0.024908414110541344
step: 280, loss: 0.002415918745100498
step: 290, loss: 0.0981629490852356
step: 300, loss: 0.011661007069051266
step: 310, loss: 0.0042647444643080235
step: 320, loss: 0.042825847864151
step: 330, loss: 0.009098992682993412
step: 340, loss: 0.007547853048890829
step: 350, loss: 0.18189483880996704
step: 360, loss: 0.006977149751037359
step: 370, loss: 0.004049696959555149
step: 380, loss: 0.024744275957345963
step: 390, loss: 0.00362958264304325
step: 400, loss: 0.0010300854919478297
step: 410, loss: 0.0038635358214378357
step: 420, loss: 0.004833395592868328
step: 430, loss: 0.01564573124051094
step: 440, loss: 0.01614040695130825
step: 450, loss: 0.07544594258069992
step: 460, loss: 0.07355823367834091
step: 470, loss: 0.00600997032597661
step: 480, loss: 0.0022476513404399157
step: 490, loss: 0.00047000456834211946
step: 500, loss: 0.035125263035297394
step: 510, loss: 0.11787791550159454
step: 520, loss: 0.035598594695329666
step: 530, loss: 0.03977608308196068
/usr/local/lib/python3.7/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=0.942634235888022, f1=0.9404706968158745,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0106655587149560452
step: 10, loss: 0.010358071886003017
step: 20, loss: 0.0012355550851561129
step: 30, loss: 0.00305117922835052
step: 40, loss: 0.00033146661007776856
step: 50, loss: 0.0938207134604454
step: 60, loss: 0.015245805494487286
step: 70, loss: 0.005163817200809717
step: 80, loss: 0.0904635488986969
```

```
step: 90, loss: 0.08651610463857651
step: 100, loss: 0.029071832075715065
step: 110, loss: 0.07019885629415512
step: 120, loss: 0.22829663753509521
step: 130, loss: 0.020894451066851616
step: 140, loss: 0.0020142709836363792
step: 150, loss: 0.008494805544614792
step: 160, loss: 0.02511458843946457
step: 170, loss: 0.008415025658905506
step: 180, loss: 0.013708122074604034
step: 190, loss: 0.015136746689677238
step: 200, loss: 0.008947491645812988
step: 210, loss: 0.001710396376438439
step: 220, loss: 0.0028408574871718884
step: 230, loss: 0.004369794391095638
step: 240, loss: 0.01127384603023529
step: 250, loss: 0.07782510668039322
step: 260, loss: 0.0012358566746115685
step: 270, loss: 0.0032426801044493914
step: 280, loss: 0.01613476499915123
step: 290, loss: 0.03508012369275093
step: 300, loss: 0.0348147377371788
step: 310, loss: 0.020718486979603767
step: 320, loss: 0.07610936462879181
step: 330, loss: 0.00048186443746089935
step: 340, loss: 0.044384364038705826
step: 350, loss: 0.0009896523552015424
step: 360, loss: 0.0022815843112766743
step: 370, loss: 0.0012491969391703606
step: 380, loss: 0.0019501617643982172
step: 390, loss: 0.155991330742836
step: 400, loss: 0.008337480016052723
step: 410, loss: 0.07881129533052444
step: 420, loss: 0.14689020812511444
step: 430, loss: 0.007172487210482359
step: 440, loss: 0.002176643116399646
step: 450, loss: 0.004746440332382917
step: 460, loss: 0.010198228061199188
step: 470, loss: 0.009053689427673817
step: 480, loss: 0.025266554206609726
step: 490, loss: 0.002325222361832857
step: 500, loss: 0.01674053631722927
step: 510, loss: 0.0015232838923111558
step: 520, loss: 0.08319573849439621
step: 530, loss: 0.0015444152522832155
/usr/local/lib/python3.7/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=0.9460853258321612, f1=0.9456928838951311,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.057392366230487823
step: 10, loss: 0.012191049754619598
step: 20, loss: 0.009517380967736244
step: 30, loss: 0.001002670032903552
step: 40, loss: 0.00026254760450683534
step: 50, loss: 0.00013956289330963045
step: 60, loss: 0.0003668347781058401
step: 70, loss: 0.0006192882428877056
step: 80, loss: 0.00025394244585186243
step: 90, loss: 0.0019304779125377536
step: 100, loss: 0.0018160765757784247
step: 110, loss: 0.0005743771907873452
step: 120, loss: 0.001033829408697784
step: 130, loss: 0.005758932325989008
step: 140, loss: 0.0011976856039837003
step: 150, loss: 0.0004822368791792542
step: 160, loss: 0.22512589395046234
step: 170, loss: 0.0018806875450536609
step: 180, loss: 0.00334237702190876
step: 190, loss: 0.029153790324926376
step: 200, loss: 0.0096417972818017
step: 210, loss: 0.0307129118591547
step: 220, loss: 0.0027423598803579807
step: 230, loss: 0.0039813038893043995
step: 240, loss: 0.006766068283468485
step: 250, loss: 0.05746425688266754
step: 260, loss: 0.0030331334564834833
step: 270, loss: 0.004462378099560738
step: 280, loss: 0.004527613054960966
step: 290, loss: 0.0020853544119745493
step: 300, loss: 0.013000193051993847
step: 310, loss: 0.049838654696941376
step: 320, loss: 0.00018409144831821322
step: 330, loss: 0.028236595913767815
step: 340, loss: 0.0005293749854899943
step: 350, loss: 0.0034504320938140154
step: 360, loss: 0.061099037528038025
step: 370, loss: 0.0015369419706985354
step: 380, loss: 0.0013169030426070094
step: 390, loss: 0.00045516956015489995
step: 400, loss: 0.0020489897578954697
step: 410, loss: 0.009429278783500195
step: 420, loss: 0.04548126831650734
```

```
step: 430, loss: 0.00017526798183098435
step: 440, loss: 0.00016075087478384376
step: 450, loss: 0.11587022989988327
step: 460, loss: 0.0007454882143065333
step: 470, loss: 0.059065088629722595
step: 480, loss: 0.015633169561624527
step: 490, loss: 0.014834129251539707
step: 500, loss: 0.003667630488052964
step: 510, loss: 0.228742316365242
step: 520, loss: 0.001647793804295361
step: 530, loss: 0.0652436763048172
/usr/local/lib/python3.7/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=0.9440591770688858, f1=0.9322191272051997,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033109881915152073
step: 10, loss: 0.0021580758038908243
step: 20, loss: 0.0010502751683816314
step: 30, loss: 0.029547642916440964
step: 40, loss: 0.005155553575605154
step: 50, loss: 0.07446962594985962
step: 60, loss: 0.0009840675629675388
step: 70, loss: 0.0006336462101899087
step: 80, loss: 0.0009767868323251605
step: 90, loss: 0.00012554405839182436
step: 100, loss: 0.03661440685391426
step: 110, loss: 0.0002850943710654974
step: 120, loss: 0.0007819727179594338
step: 130, loss: 0.00021137064322829247
step: 140, loss: 0.0007869385881349444
step: 150, loss: 0.005334565881639719
step: 160, loss: 0.00030191574478521943
step: 170, loss: 0.0017309539252892137
step: 180, loss: 0.017357490956783295
step: 190, loss: 0.005358547437936068
step: 200, loss: 0.005061317700892687
step: 210, loss: 0.005527989473193884
step: 220, loss: 0.0019543091766536236
step: 230, loss: 0.0006069173687137663
step: 240, loss: 0.0006552666891366243
step: 250, loss: 0.004187586717307568
step: 260, loss: 0.0012421113206073642
step: 270, loss: 0.00026862454251386225
step: 280, loss: 0.0006542777991853654
```

```
step: 290, loss: 0.0007062662625685334
step: 300, loss: 0.00041665358003228903
step: 310, loss: 0.0002244054339826107
step: 320, loss: 0.057284969836473465
step: 330, loss: 0.0276962798088789
step: 340, loss: 0.0008890336030162871
step: 350, loss: 0.00037508236709982157
step: 360, loss: 0.0038711782544851303
step: 370, loss: 0.024402985349297523
step: 380, loss: 0.0058473385870456696
step: 390, loss: 0.004578061867505312
step: 400, loss: 0.0048031373880803585
step: 410, loss: 0.0002382095844950527
step: 420, loss: 0.03018653206527233
step: 430, loss: 0.00016619425150565803
step: 440, loss: 0.0010359379230067134
step: 450, loss: 0.012407107278704643
step: 460, loss: 0.048354413360357285
step: 470, loss: 0.0523470900952816
step: 480, loss: 0.006377441808581352
step: 490, loss: 0.008465910330414772
step: 500, loss: 0.007859979756176472
step: 510, loss: 0.028908250853419304
step: 520, loss: 0.036303043365478516
step: 530, loss: 0.0038966743741184473
/usr/local/lib/python3.7/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=0.9461862423958821, f1=0.940354147250699,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005677745793946087
step: 10, loss: 0.00268009421415627
step: 20, loss: 0.0018966739298775792
step: 30, loss: 0.0008639958105050027
step: 40, loss: 0.001059932284988463
step: 50, loss: 0.00014909979654476047
step: 60, loss: 0.0005247521330602467
step: 70, loss: 0.0014534194488078356
step: 80, loss: 0.00032436667243018746
step: 90, loss: 0.00035905250115320086
step: 100, loss: 0.0005633653490804136
step: 110, loss: 0.0005303557845763862
step: 120, loss: 0.009918199852108955
step: 130, loss: 0.0008542143041267991
step: 140, loss: 0.023215344175696373
```

```
step: 150, loss: 0.0006252937600947917
step: 160, loss: 0.0002529580087866634
step: 170, loss: 0.006061016581952572
step: 180, loss: 0.0002658884914126247
step: 190, loss: 0.0025313349906355143
step: 200, loss: 0.0007728647906333208
step: 210, loss: 0.10886228084564209
step: 220, loss: 9.976655564969406e-05
step: 230, loss: 0.060293618589639664
step: 240, loss: 0.01704830676317215
step: 250, loss: 0.0032751804683357477
step: 260, loss: 0.0002737132890615612
step: 270, loss: 0.022706270217895508
step: 280, loss: 0.01266727689653635
step: 290, loss: 0.0001604948192834854
step: 300, loss: 0.0001189977556350641
step: 310, loss: 0.0008152697701007128
step: 320, loss: 0.0005677470471709967
step: 330, loss: 0.0009113112464547157
step: 340, loss: 0.01621050015091896
step: 350, loss: 0.0003390965866856277
step: 360, loss: 0.012070106342434883
step: 370, loss: 0.00024449435295537114
step: 380, loss: 0.00012195105955470353
step: 390, loss: 0.0024203697685152292
step: 400, loss: 0.00011150696082040668
step: 410, loss: 0.001367596909403801
step: 420, loss: 0.00010713381925597787
step: 430, loss: 0.017371494323015213
step: 440, loss: 0.0252092182636261
step: 450, loss: 0.004946931265294552
step: 460, loss: 0.0006672806921415031
step: 470, loss: 0.028594255447387695
step: 480, loss: 0.006246991455554962
step: 490, loss: 0.011914962902665138
step: 500, loss: 0.0005177879356779158
step: 510, loss: 0.00020861451048403978
step: 520, loss: 0.00015260757936630398
step: 530, loss: 0.05755957216024399
/usr/local/lib/python3.7/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=0.9473684210526315, f1=0.9380281690140844,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011471015022834763
```

```
step: 10, loss: 0.08265269547700882
step: 20, loss: 0.015245708636939526
step: 30, loss: 0.05299757048487663
step: 40, loss: 0.001162671484053135
step: 50, loss: 0.0015678835334256291
step: 60, loss: 0.00025959560298360884
step: 70, loss: 0.0009726957650855184
step: 80, loss: 0.0013439549366012216
step: 90, loss: 0.011126021854579449
step: 100, loss: 0.00023851983132772148
step: 110, loss: 0.00718596251681447
step: 120, loss: 8.449427696177736e-05
step: 130, loss: 0.00016441625484731048
step: 140, loss: 0.00017797162581700832
step: 150, loss: 0.009018021635711193
step: 160, loss: 0.0009090102976188064
step: 170, loss: 0.0006238378118723631
step: 180, loss: 0.16118139028549194
step: 190, loss: 0.001387608703225851
step: 200, loss: 0.0006485198391601443
step: 210, loss: 0.0027927502524107695
step: 220, loss: 0.003973150160163641
step: 230, loss: 0.0011767406249418855
step: 240, loss: 0.0004891353310085833
step: 250, loss: 0.0004795287095475942
step: 260, loss: 0.0007711007492616773
step: 270, loss: 0.00031657537329010665
step: 280, loss: 0.007055401802062988
step: 290, loss: 0.0011316542513668537
step: 300, loss: 0.002075174590572715
step: 310, loss: 0.001562587101943791
step: 320, loss: 0.00019316354882903397
step: 330, loss: 0.0010276301763951778
step: 340, loss: 0.0005498608225025237
step: 350, loss: 0.02065635472536087
step: 360, loss: 0.01836244761943817
step: 370, loss: 8.109274494927377e-05
step: 380, loss: 0.0030476334504783154
step: 390, loss: 5.497169331647456e-05
step: 400, loss: 0.004466277081519365
step: 410, loss: 0.00040261601679958403
step: 420, loss: 0.0028981186915189028
step: 430, loss: 0.0024183164350688457
step: 440, loss: 9.761922410689294e-05
step: 450, loss: 0.0017562095308676362
step: 460, loss: 0.016564853489398956
step: 470, loss: 8.046960283536464e-05
step: 480, loss: 0.003595230635255575
```

```
step: 490, loss: 0.06830314546823502
step: 500, loss: 0.016704322770237923
step: 510, loss: 0.011267991736531258
step: 520, loss: 0.0560898520052433
step: 530, loss: 0.03480195626616478
/usr/local/lib/python3.7/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=0.9509259259259258, f1=0.9392675011590171,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005106363678351045
step: 10, loss: 0.00016351402155512365
step: 20, loss: 0.0003634168242570013
step: 30, loss: 0.009921331889927387
step: 40, loss: 0.0027752236928790808
step: 50, loss: 0.00022340404393617064
step: 60, loss: 0.007330005988478661
step: 70, loss: 8.305454684887081e-05
step: 80, loss: 6.391304486896843e-05
step: 90, loss: 5.7210119848605245e-05
step: 100, loss: 0.00037678805529139936
step: 110, loss: 0.00013172056060284376
step: 120, loss: 0.00020944415882695466
step: 130, loss: 0.008386199362576008
step: 140, loss: 0.0003631678409874439
step: 150, loss: 7.619855023222044e-05
step: 160, loss: 0.018572669476270676
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 5.597104245680384e-05
step: 180, loss: 0.003425830975174904
step: 190, loss: 6.247318378882483e-05
step: 200, loss: 5.550163041334599e-05
step: 210, loss: 0.007644155062735081
step: 220, loss: 0.00011103774886578321
step: 230, loss: 0.001918287482112646
step: 240, loss: 0.00010304132592864335
step: 250, loss: 5.405855335993692e-05
step: 260, loss: 0.004953599534928799
step: 270, loss: 8.013199840206653e-05
step: 280, loss: 0.015852224081754684
step: 290, loss: 0.0007414116407744586
step: 300, loss: 0.0011574913514778018
step: 310, loss: 0.04762653261423111
step: 320, loss: 0.09816916286945343
step: 330, loss: 0.0013497439213097095
```

```
step: 340, loss: 0.0002766909310594201
step: 350, loss: 0.0004117066855542362
step: 360, loss: 0.0001059797577909194
step: 370, loss: 0.00013943779049441218
step: 380, loss: 0.007011003792285919
step: 390, loss: 0.0046132756397128105
step: 400, loss: 0.00030556487035937607
step: 410, loss: 0.005317705217774888
step: 420, loss: 0.00014062653644941747
step: 430, loss: 0.0006798029062338173
step: 440, loss: 0.0003109807730652392
step: 450, loss: 0.010836598463356495
step: 460, loss: 0.002797291614115238
step: 470, loss: 0.005407918244600296
step: 480, loss: 0.00013467279495671391
step: 490, loss: 0.003297598799690604
step: 500, loss: 0.003909296356141567
step: 510, loss: 7.652933709323406e-05
step: 520, loss: 7.910842396086082e-05
step: 530, loss: 0.00014207397180143744
/usr/local/lib/python3.7/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=0.9492003762935088, f1=0.9380863039399624,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.909344276413321e-05
step: 10, loss: 0.0013688185717910528
step: 20, loss: 0.003507968969643116
step: 30, loss: 0.001650824910029769
step: 40, loss: 0.002764287404716015
step: 50, loss: 5.601411612587981e-05
step: 60, loss: 0.0022794490214437246
step: 70, loss: 0.0005873314221389592
step: 80, loss: 0.0016908016987144947
step: 90, loss: 8.79660583450459e-05
step: 100, loss: 4.6632438170490786e-05
step: 110, loss: 0.0001136002319981344
step: 120, loss: 0.00010188585292780772
step: 130, loss: 7.174373604357243e-05
step: 140, loss: 5.932354542892426e-05
step: 150, loss: 4.732287197839469e-05
step: 160, loss: 2.445532300043851e-05
step: 170, loss: 3.706593270180747e-05
step: 180, loss: 0.00010253386426484212
step: 190, loss: 0.01820121519267559
```

```
step: 200, loss: 0.0008487799786962569
step: 210, loss: 9.740863606566563e-05
step: 220, loss: 0.015470059588551521
step: 230, loss: 8.169277862180024e-05
step: 240, loss: 0.0018870129715651274
step: 250, loss: 0.00011884517880389467
step: 260, loss: 0.0007142442627809942
step: 270, loss: 0.0004277661209926009
step: 280, loss: 0.000707845261786133
step: 290, loss: 0.0035058981738984585
step: 300, loss: 4.431296474649571e-05
step: 310, loss: 0.0004608076123986393
step: 320, loss: 0.0011457785731181502
step: 330, loss: 4.0293962229043245e-05
step: 340, loss: 7.02625184203498e-05
step: 350, loss: 0.00034581052022986114
step: 360, loss: 0.0014079725369811058
step: 370, loss: 0.00045845116255804896
step: 380, loss: 8.390426228288561e-05
step: 390, loss: 2.9249391445773654e-05
step: 400, loss: 0.00012549843813758343
step: 410, loss: 0.001744030392728746
step: 420, loss: 6.37484117760323e-05
step: 430, loss: 0.002399513265118003
step: 440, loss: 0.00010998837387887761
step: 450, loss: 8.965947199612856e-05
step: 460, loss: 8.638412691652775e-05
step: 470, loss: 0.00014102619024924934
step: 480, loss: 8.417649223702028e-05
step: 490, loss: 7.501107756979764e-05
step: 500, loss: 8.122285362333059e-05
step: 510, loss: 0.003438031766563654
step: 520, loss: 0.00012165111547801644
step: 530, loss: 0.0001896568137453869
/usr/local/lib/python3.7/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=0.9489461358313818, f1=0.9394221808014911,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.74800071970094e-05
step: 10, loss: 0.00017843143723439425
step: 20, loss: 0.0002500055416021496
step: 30, loss: 5.5266922572627664e-05
step: 40, loss: 0.00011547709436854348
step: 50, loss: 0.00012486876221373677
```

```
step: 60, loss: 8.713408897165209e-05
step: 70, loss: 6.133350689196959e-05
step: 80, loss: 0.2167758047580719
step: 90, loss: 0.0001820926700020209
step: 100, loss: 0.011282875202596188
step: 110, loss: 0.00023509032325819135
step: 120, loss: 0.00020642300660256296
step: 130, loss: 0.0001924337266245857
step: 140, loss: 0.00020285944628994912
step: 150, loss: 0.00015483246534131467
step: 160, loss: 9.972054249374196e-05
step: 170, loss: 0.0001795903081074357
step: 180, loss: 0.0001909438578877598
step: 190, loss: 0.0004431271518114954
step: 200, loss: 3.529758396325633e-05
step: 210, loss: 0.00013934042362961918
step: 220, loss: 8.834482287056744e-05
step: 230, loss: 0.00015525588241871446
step: 240, loss: 0.0001417262537870556
step: 250, loss: 6.871263030916452e-05
step: 260, loss: 9.038708958541974e-05
step: 270, loss: 0.0005381760420277715
step: 280, loss: 0.00010647252202033997
step: 290, loss: 0.00010900241613853723
step: 300, loss: 0.00020876912458334118
step: 310, loss: 0.0001272650551982224
step: 320, loss: 8.514778164681047e-05
step: 330, loss: 0.0002024590503424406
step: 340, loss: 4.051774158142507e-05
step: 350, loss: 2.655202661117073e-05
step: 360, loss: 7.361797906924039e-05
step: 370, loss: 8.619485015515238e-05
step: 380, loss: 0.0030497952830046415
step: 390, loss: 0.0004567424184642732
step: 400, loss: 6.410532660083845e-05
step: 410, loss: 0.00011820586951216683
step: 420, loss: 0.0034826924093067646
step: 430, loss: 7.592875772388652e-05
step: 440, loss: 7.003189966781065e-05
step: 450, loss: 0.00010512759763514623
step: 460, loss: 4.874672231380828e-05
step: 470, loss: 0.0025168282445520163
step: 480, loss: 5.5764954595360905e-05
step: 490, loss: 5.517998579307459e-05
step: 500, loss: 3.207149347872473e-05
step: 510, loss: 5.263035200187005e-05
step: 520, loss: 7.388075755443424e-05
step: 530, loss: 0.000337899720761925
```

```
/usr/local/lib/python3.7/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=0.9421028253821214, f1=0.9353647276084949,
best_f1=0.9429234338747099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002818310749717057
step: 10, loss: 0.00012423649604897946
step: 20, loss: 0.005254008807241917
step: 30, loss: 0.00049735454376786 95
step: 40, loss: 0.00020532285270746797
step: 50, loss: 0.005278981290757656
step: 60, loss: 0.0006119439494796097
step: 70, loss: 0.0002004775742534548
step: 80, loss: 0.001011675805784762
step: 90, loss: 0.00044696693657897413
step: 100, loss: 0.0002893750206567347
step: 110, loss: 6.510586536023766e-05
step: 120, loss: 0.001438617124222219
step: 130, loss: 0.0003514715644996613
step: 140, loss: 6.709950685035437e-05
step: 150, loss: 0.00014173590170685202
step: 160, loss: 0.00014761323109269142
step: 170, loss: 0.00014213712711352855
step: 180, loss: 0.00010936409671558067
step: 190, loss: 0.00028268288588151336
step: 200, loss: 0.00029269259539432824
step: 210, loss: 0.00010801203461596742
step: 220, loss: 0.00031955799204297364
step: 230, loss: 0.006727786269038916
step: 240, loss: 0.0005723318899981678
step: 250, loss: 0.0005225933855399489
step: 260, loss: 0.000149071347550489
step: 270, loss: 0.00014498854579869658
step: 280, loss: 0.008425370790064335
step: 290, loss: 3.20762483170256e-05
step: 300, loss: 6.11240029684268e-05
step: 310, loss: 0.0002691211993806064
step: 320, loss: 0.0010297447443008423
step: 330, loss: 0.0002516719396226108
step: 340, loss: 0.0008685330394655466
step: 350, loss: 0.00013453964493237436
step: 360, loss: 0.00026397337205708027
step: 370, loss: 0.003992526326328516
step: 380, loss: 0.0001122230023611337
step: 390, loss: 8.351581345777959e-05
```

```
step: 400, loss: 0.0008971744682639837
step: 410, loss: 0.000343810097547248
step: 420, loss: 0.0001781676255632192
step: 430, loss: 0.00022394808183889836
step: 440, loss: 0.00011063116107834503
step: 450, loss: 0.006345644593238831
step: 460, loss: 0.00017557803948875517
step: 470, loss: 0.0030354647897183895
step: 480, loss: 4.405408617458306e-05
step: 490, loss: 6.637969636358321e-05
step: 500, loss: 0.00069381570210680037
step: 510, loss: 7.156419451348484e-05
step: 520, loss: 0.0001407470554113388
step: 530, loss: 6.011690857121721e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9526901669758813, f1=0.9405255878284925,
best_f1=0.9405255878284925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.820317791309208e-05
step: 10, loss: 0.00012751336907967925
step: 20, loss: 6.053956167306751e-05
step: 30, loss: 0.00016016527661122382
step: 40, loss: 7.799240847816691e-05
step: 50, loss: 0.0004670399648603052
step: 60, loss: 0.0013593211770057678
step: 70, loss: 0.0001403878559358418
step: 80, loss: 0.00010204065620200709
step: 90, loss: 5.992900696583092e-05
step: 100, loss: 0.0021844827570021152
step: 110, loss: 0.004516819957643747
step: 120, loss: 8.240262832259759e-05
step: 130, loss: 4.6238379582064226e-05
step: 140, loss: 6.994631257839501e-05
step: 150, loss: 0.00014873083273414522
step: 160, loss: 0.000676741125062108
step: 170, loss: 8.110165072139353e-05
step: 180, loss: 4.244543379172683e-05
step: 190, loss: 5.8173340221401304e-05
step: 200, loss: 0.00013715167006012052
step: 210, loss: 8.697580778971314e-05
```

```
step: 220, loss: 4.4590655306819826e-05
step: 230, loss: 0.0001110700613935478
step: 240, loss: 3.0942745070206e-05
step: 250, loss: 0.0007311474182642996
step: 260, loss: 4.978665674570948e-05
step: 270, loss: 6.828927871538326e-05
step: 280, loss: 5.8213543525198475e-05
step: 290, loss: 0.00012036525731673464
step: 300, loss: 7.719513814663514e-05
step: 310, loss: 8.208152576116845e-05
step: 320, loss: 6.573555583599955e-05
step: 330, loss: 0.0002596111153252423
step: 340, loss: 0.00010269045014865696
step: 350, loss: 3.4013602999038994e-05
step: 360, loss: 0.00011948532483074814
step: 370, loss: 4.424104918143712e-05
step: 380, loss: 0.0007949264836497605
step: 390, loss: 3.7706136936321855e-05
step: 400, loss: 0.004584718961268663
step: 410, loss: 3.718443258549087e-05
step: 420, loss: 0.00030167106888256967
step: 430, loss: 0.0003794893273152411
step: 440, loss: 0.00040385028114542365
step: 450, loss: 6.262245733523741e-05
step: 460, loss: 0.006357175763696432
step: 470, loss: 0.0003050586674362421
step: 480, loss: 6.193656008690596e-05
step: 490, loss: 0.00025523401563987136
step: 500, loss: 6.069578739698045e-05
step: 510, loss: 0.0034195231273770332
step: 520, loss: 3.5230379580752924e-05
step: 530, loss: 8.250109385699034e-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=0.9497927222478121, f1=0.939990838295923,
best_f1=0.9405255878284925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.8643072912236676e-05
step: 10, loss: 9.577375749358907e-05
step: 20, loss: 0.00014615949476137757
step: 40, loss: 5.317376781022176e-05
step: 50, loss: 8.213447290472686e-05
step: 60, loss: 9.283188410336152e-05
step: 70, loss: 7.318323332583532e-05
step: 80, loss: 3.814432056969963e-05
```

```
step: 90, loss: 7.862594065954909e-05
step: 100, loss: 1.7868991562863812e-05
step: 110, loss: 0.0001114753758884035
step: 120, loss: 8.726974192541093e-05
step: 130, loss: 3.931383616873063e-05
step: 140, loss: 7.901558274170384e-05
step: 150, loss: 7.915832247817889e-05
step: 160, loss: 8.568317571189255e-05
step: 170, loss: 0.00011434023326728493
step: 180, loss: 7.830396498320624e-05
step: 190, loss: 7.64240903663449e-05
step: 200, loss: 0.0001709428761387244
step: 210, loss: 9.140741167357191e-05
step: 220, loss: 3.682955866679549e-05
step: 230, loss: 0.22759667038917542
step: 240, loss: 6.277101783780381e-05
step: 250, loss: 4.213371721561998e-05
step: 260, loss: 5.892973786103539e-05
step: 270, loss: 0.0002484588767401874
step: 280, loss: 4.9402959120925516e-05
step: 290, loss: 9.240769577445462e-05
step: 300, loss: 4.397556040203199e-05
step: 310, loss: 0.005060078576207161
step: 320, loss: 8.478185918647796e-05
step: 330, loss: 7.63748976169154e-05
step: 340, loss: 0.0002490198239684105
step: 350, loss: 0.0060022673569619656
step: 360, loss: 0.0005017844960093498
step: 370, loss: 0.0005424338160082698
step: 380, loss: 0.00016395573038607836
step: 390, loss: 5.073286592960358e-05
step: 400, loss: 0.0067386026494205
step: 410, loss: 5.840449739480391e-05
step: 420, loss: 3.362172355991788e-05
step: 430, loss: 2.2908814571565017e-05
step: 440, loss: 3.403138180146925e-05
step: 450, loss: 0.004492718260735273
step: 460, loss: 6.789770850446075e-05
step: 470, loss: 4.080664075445384e-05
step: 480, loss: 4.9866244808072224e-05
step: 490, loss: 6.626891263294965e-05
step: 500, loss: 0.000221827402128838
step: 510, loss: 7.17878938303329e-05
step: 520, loss: 8.206891652662307e-05
step: 530, loss: 7.340708543779328e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
epoch 15: dev_f1=0.9506517690875234, f1=0.9453197405004634,
best_f1=0.9405255878284925
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, 225.38it/s]
load_f1 = 0.952292728114868
real_f1 = 0.9517177344475395
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, 184.59it/s]
```

2.1.4 Amazon-Google - Running the matcher

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

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

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

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

Selected optimization level 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: 0.4981679320335388
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5173988938331604
step: 20, loss: 0.5504347085952759
step: 30, loss: 0.3539438247680664
step: 40, loss: 0.3436773419380188
step: 50, loss: 0.45509007573127747
step: 60, loss: 0.46238210797309875
step: 70, loss: 0.2889605462551117
step: 80, loss: 0.3236367106437683
step: 90, loss: 0.2432655245065689
step: 100, loss: 0.30678075551986694
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.24831053614616394
step: 120, loss: 0.3644711375236511
step: 130, loss: 0.2621496319770813
step: 140, loss: 0.3618651330471039
step: 150, loss: 0.22438593208789825
step: 160, loss: 0.3994426131248474
step: 170, loss: 0.20719943940639496
step: 180, loss: 0.269420862197876
step: 190, loss: 0.40080639719963074
step: 200, loss: 0.21310296654701233
step: 210, loss: 0.3566439151763916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.463302752293578, f1=0.4963503649635037,
best_f1=0.4963503649635037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.229786217212677
step: 10, loss: 0.07689961791038513
step: 20, loss: 0.3106960356235504
step: 30, loss: 0.18895968794822693
step: 40, loss: 0.4556652009487152
step: 50, loss: 0.18155671656131744
step: 60, loss: 0.27136391401290894
step: 70, loss: 0.15913915634155273
step: 80, loss: 0.20057258009910583
step: 90, loss: 0.14770960807800293
step: 100, loss: 0.3299799859523773
step: 110, loss: 0.29178386926651
step: 120, loss: 0.1503559798002243
step: 130, loss: 0.22960415482521057
step: 140, loss: 0.0865735113620758
step: 150, loss: 0.2517207860946655
step: 160, loss: 0.13766244053840637
step: 170, loss: 0.3006531596183777
step: 180, loss: 0.19738751649856567
step: 190, loss: 0.32296258211135864
step: 200, loss: 0.0917648896574974
step: 210, loss: 0.17158332467079163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5836734693877551, f1=0.5720338983050847,
best_f1=0.5720338983050847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0692952424287796
step: 10, loss: 0.07464870065450668
step: 20, loss: 0.30386635661125183
step: 30, loss: 0.08401407301425934
step: 40, loss: 0.21959364414215088
step: 50, loss: 0.13696713745594025
step: 60, loss: 0.30729061365127563
step: 70, loss: 0.08041137456893921
step: 80, loss: 0.1562788337469101
step: 90, loss: 0.21875715255737305
step: 100, loss: 0.2580798268318176
step: 110, loss: 0.04495464265346527
step: 120, loss: 0.169867143034935
```

```
step: 130, loss: 0.0850544199347496
step: 140, loss: 0.1190018355846405
step: 150, loss: 0.11267660558223724
step: 160, loss: 0.09126506000757217
step: 170, loss: 0.1653921902179718
step: 180, loss: 0.04679998755455017
step: 190, loss: 0.023702332749962807
step: 200, loss: 0.07079646736383438
step: 210, loss: 0.20668232440948486
/usr/local/lib/python3.7/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=0.5805168986083499, f1=0.5874999999999999,
best_f1=0.5720338983050847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09373167902231216
step: 10, loss: 0.1410255879163742
step: 20, loss: 0.06857169419527054
step: 30, loss: 0.1098635122179985
step: 40, loss: 0.06637706607580185
step: 50, loss: 0.2143716663122177
step: 60, loss: 0.11506254971027374
step: 70, loss: 0.02319597825407982
step: 80, loss: 0.09948400408029556
step: 90, loss: 0.09291846305131912
step: 100, loss: 0.14959552884101868
step: 110, loss: 0.4482012689113617
step: 120, loss: 0.12937146425247192
step: 130, loss: 0.19172590970993042
step: 140, loss: 0.40773123502731323
step: 150, loss: 0.08616004884243011
step: 160, loss: 0.27213120460510254
step: 170, loss: 0.08214868605136871
step: 180, loss: 0.08460266888141632
step: 190, loss: 0.09796422719955444
step: 200, loss: 0.07821100205183029
step: 210, loss: 0.361674040555954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6135458167330677, f1=0.6378600823045267,
best_f1=0.6378600823045267
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.178626120090482462
step: 10, loss: 0.0705099031329155
step: 20, loss: 0.25592634081840515
step: 30, loss: 0.010942311957478523
step: 40, loss: 0.0993170216679573
step: 50, loss: 0.10137154906988144
step: 60, loss: 0.09266039729118347
step: 70, loss: 0.15463300049304962
step: 80, loss: 0.0764077827334404
step: 90, loss: 0.036857761442661285
step: 100, loss: 0.029933908954262733
step: 110, loss: 0.06178809702396393
step: 120, loss: 0.1271403729915619
step: 130, loss: 0.08609358966350555
step: 140, loss: 0.20304691791534424
step: 150, loss: 0.07779614627361298
step: 160, loss: 0.07026709616184235
step: 170, loss: 0.08331198245286942
step: 180, loss: 0.13508816063404083
step: 190, loss: 0.14497263729572296
step: 200, loss: 0.16942888498306274
step: 210, loss: 0.1200256496667862
/usr/local/lib/python3.7/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=0.610752688172043, f1=0.5823927765237021,
best_f1=0.6378600823045267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02770555391907692
step: 10, loss: 0.2145388424396515
step: 20, loss: 0.010101447813212872
step: 30, loss: 0.070191890001297
step: 40, loss: 0.06598759442567825
step: 50, loss: 0.19253185391426086
step: 60, loss: 0.09313416481018066
step: 70, loss: 0.07525275647640228
step: 80, loss: 0.0865064263343811
step: 90, loss: 0.08893046528100967
step: 100, loss: 0.23440459370613098
step: 110, loss: 0.05907895416021347
step: 120, loss: 0.007010876666754484
step: 130, loss: 0.07222636789083481
step: 140, loss: 0.11254173517227173
```

```
step: 150, loss: 0.06924577802419662
step: 160, loss: 0.11987002193927765
step: 170, loss: 0.10171417146921158
step: 180, loss: 0.0617542527616024
step: 190, loss: 0.025292007252573967
step: 200, loss: 0.09942233562469482
step: 210, loss: 0.11962857097387314
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6274509803921569, f1=0.6055045871559633,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030591877177357674
step: 10, loss: 0.03414880484342575
step: 20, loss: 0.016934430226683617
step: 30, loss: 0.059012122452259064
step: 40, loss: 0.08935277163982391
step: 50, loss: 0.018044402822852135
step: 60, loss: 0.017978226765990257
step: 70, loss: 0.06710315495729446
step: 80, loss: 0.03594575822353363
step: 90, loss: 0.03328325226902962
step: 100, loss: 0.06768199056386948
step: 110, loss: 0.0426567867398262
step: 120, loss: 0.02403968945145607
step: 130, loss: 0.2741098403930664
step: 140, loss: 0.02133958786725998
step: 150, loss: 0.09682206064462662
step: 160, loss: 0.14861157536506653
step: 170, loss: 0.04294706508517265
step: 180, loss: 0.01008442509919405
step: 190, loss: 0.10416161268949509
step: 200, loss: 0.03545968234539032
step: 210, loss: 0.15214703977108002
/usr/local/lib/python3.7/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=0.6243194192377496, f1=0.6048237476808905,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11271661520004272
step: 10, loss: 0.028728144243359566
step: 20, loss: 0.013890138827264309
step: 30, loss: 0.0525507889688015
step: 40, loss: 0.018349653109908104
step: 50, loss: 0.0009034443646669388
step: 60, loss: 0.07196105271577835
step: 70, loss: 0.08639732748270035
step: 80, loss: 0.10791893303394318
step: 90, loss: 0.07857102155685425
step: 100, loss: 0.11128589510917664
step: 110, loss: 0.009659739211201668
step: 120, loss: 0.13694584369659424
step: 130, loss: 0.00021189922699704766
step: 140, loss: 0.006512507330626249
step: 150, loss: 0.20953300595283508
step: 160, loss: 0.0826481357216835
step: 170, loss: 0.03772638365626335
step: 180, loss: 0.013456437736749649
step: 190, loss: 0.005548458080738783
step: 200, loss: 0.04585276171565056
step: 210, loss: 0.12451866269111633
/usr/local/lib/python3.7/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=0.6242990654205607, f1=0.6179159049360147,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021023042500019073
step: 10, loss: 0.010758372955024242
step: 20, loss: 0.07548482716083527
step: 30, loss: 0.012160050682723522
step: 40, loss: 0.001555610098876059
step: 50, loss: 0.025290317833423615
step: 60, loss: 0.05990368500351906
step: 70, loss: 0.006621664855629206
step: 80, loss: 0.00174332142341882
step: 90, loss: 0.0029377511236816645
step: 100, loss: 0.012866931036114693
step: 110, loss: 0.0688488632440567
step: 120, loss: 0.08075393736362457
step: 130, loss: 0.002058525336906314
step: 140, loss: 0.05534975230693817
step: 150, loss: 0.0012380112893879414
step: 160, loss: 0.02476189099252224
```

```
step: 170, loss: 0.005489376373589039
step: 180, loss: 0.036284055560827255
step: 190, loss: 0.005901539698243141
step: 200, loss: 0.003975085914134979
step: 210, loss: 0.009852472692728043
/usr/local/lib/python3.7/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=0.6102362204724409, f1=0.6000000000000001,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030925471801310778
step: 10, loss: 0.04293908551335335
step: 20, loss: 0.006790489889681339
step: 30, loss: 0.0007896106108091772
step: 40, loss: 0.0028945652302354574
step: 50, loss: 0.061308130621910095
step: 60, loss: 0.0013960228534415364
step: 70, loss: 0.027962790802121162
step: 80, loss: 0.04958229884505272
step: 90, loss: 0.008244096301496029
step: 100, loss: 0.010920283384621143
step: 110, loss: 0.034366440027952194
step: 120, loss: 0.2991398870944977
step: 130, loss: 0.044935815036296844
step: 140, loss: 0.0031761156860738993
step: 150, loss: 0.011607528664171696
step: 160, loss: 0.024264151230454445
step: 170, loss: 0.0007039176416583359
step: 180, loss: 0.0156322680413723
step: 190, loss: 0.007344119716435671
step: 200, loss: 0.004819219466298819
step: 210, loss: 0.002181708812713623
/usr/local/lib/python3.7/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=0.6137787056367432, f1=0.597938144329897,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07243772596120834
step: 10, loss: 0.0033376680221408606
step: 20, loss: 0.001612398773431778
step: 30, loss: 0.007113362662494183
step: 40, loss: 0.002574267564341426
```

```
step: 50, loss: 0.10748410224914551
step: 60, loss: 0.037570588290691376
step: 70, loss: 0.0032454694155603647
step: 80, loss: 0.01691460609436035
step: 90, loss: 0.08697476983070374
step: 100, loss: 0.13868172466754913
step: 110, loss: 0.029651634395122528
step: 120, loss: 0.006353230215609074
step: 130, loss: 0.0028970176354050636
step: 140, loss: 0.015226038172841072
step: 150, loss: 0.003834057366475463
step: 160, loss: 0.001955899944288492
step: 170, loss: 0.009774237871170044
step: 180, loss: 0.0005864640115760267
step: 190, loss: 0.06895565241575241
step: 200, loss: 0.0008515417575836182
step: 210, loss: 0.006691932212561369
/usr/local/lib/python3.7/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=0.623015873015873, f1=0.630952380952381,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015133172273635864
step: 10, loss: 0.003474696772173047
step: 20, loss: 0.04565058648586273
step: 30, loss: 0.009966160170733929
step: 40, loss: 0.0006946671637706459
step: 50, loss: 0.006167408544570208
step: 60, loss: 0.0009084266494402827
step: 70, loss: 0.004091157577931881
step: 80, loss: 0.021837256848812103
step: 90, loss: 0.08379758149385452
step: 100, loss: 0.0004397413576953113
step: 110, loss: 0.020892782136797905
step: 120, loss: 0.0004527427372522652
step: 130, loss: 0.004605090711265802
step: 140, loss: 0.06368464976549149
step: 150, loss: 0.00012759052333422005
step: 160, loss: 0.0016492418944835663
step: 170, loss: 0.08997219055891037
step: 180, loss: 0.007704677060246468
step: 190, loss: 0.02944502793252468
step: 200, loss: 0.0003880462027154863
step: 210, loss: 0.0033721118234097958
/usr/local/lib/python3.7/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=0.6140724946695096, f1=0.5817409766454353,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020056811627000057
step: 10, loss: 0.02991967648267746
step: 20, loss: 0.0009675328619778156
step: 30, loss: 0.0008023551781661808
step: 40, loss: 0.010497777722775936
step: 50, loss: 0.04947545751929283
step: 60, loss: 0.03972368314862251
step: 70, loss: 0.009727277792990208
step: 80, loss: 0.02424766682088375
step: 90, loss: 0.006551133934408426
step: 100, loss: 0.003960023168474436
step: 110, loss: 0.026290327310562134
step: 120, loss: 0.03253438323736191
step: 130, loss: 0.00028707069577649236
step: 140, loss: 0.0012044825125485659
step: 150, loss: 0.0015270576113834977
step: 160, loss: 0.0022944340016692877
step: 170, loss: 0.0043155476450920105
step: 180, loss: 0.000343859865097329
step: 190, loss: 0.001918518915772438
step: 200, loss: 0.021503295749425888
step: 210, loss: 0.0022577897179871798
/usr/local/lib/python3.7/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=0.6153846153846153, f1=0.6208251473477407,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001766976318322122
step: 10, loss: 0.01831710711121559
step: 20, loss: 0.00047193083446472883
step: 30, loss: 0.0003679332439787686
step: 40, loss: 0.062185268849134445
step: 50, loss: 0.004250775091350079
step: 60, loss: 0.13737726211547852
step: 70, loss: 0.0007326716440729797
step: 80, loss: 0.024303598329424858
step: 90, loss: 0.00022288948821369559
step: 100, loss: 0.007338732481002808
```

```
step: 110, loss: 0.001941764377988875
step: 120, loss: 0.00043909394298680127
step: 130, loss: 0.20347769558429718
step: 140, loss: 0.0011535065714269876
step: 150, loss: 0.007721024565398693
step: 160, loss: 0.00037252847687341273
step: 170, loss: 0.005133095663040876
step: 180, loss: 0.07688067108392715
step: 190, loss: 0.07195910066366196
step: 200, loss: 0.04534371569752693
step: 210, loss: 0.0014771115966141224
/usr/local/lib/python3.7/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=0.5987261146496815, f1=0.6090712742980561,
best_f1=0.6055045871559633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0103563731536269919
step: 10, loss: 0.0008179342257790267
step: 20, loss: 0.004068182781338692
step: 30, loss: 0.0054215677082538605
step: 40, loss: 0.013856912963092327
step: 50, loss: 0.0021670283749699593
step: 60, loss: 0.0030166851356625557
step: 70, loss: 0.09849441796541214
step: 80, loss: 0.001251128502190113
step: 90, loss: 0.0008831470040604472
step: 100, loss: 0.0005182541208341718
step: 110, loss: 0.05453551933169365
step: 120, loss: 0.004084537737071514
step: 130, loss: 0.0004577215586323291
step: 140, loss: 0.015591183677315712
step: 150, loss: 0.007436505984514952
step: 160, loss: 0.0018664170056581497
step: 170, loss: 0.01278591062873602
step: 180, loss: 0.009689945727586746
step: 190, loss: 0.001346259145066142
step: 200, loss: 0.0028396742418408394
step: 210, loss: 0.009920001029968262
/usr/local/lib/python3.7/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=0.5948275862068966, f1=0.6043478260869565,
best_f1=0.6055045871559633
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 [00:06, 377.88it/s]
load_f1 = 0.6304347826086957
real_f1 = 0.6247288503253796
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 178.16it/s]
```

2.1.5 Walmart-Amazon - Running the matcher

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

```
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model

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

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

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

Defaults for this optimization level are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`.  Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4557383358478546
step: 10, loss: 0.45567086338996887
step: 20, loss: 0.24792060256004333
step: 30, loss: 0.3647560179233551
step: 40, loss: 0.1792753040790558
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.3022686839103699
step: 60, loss: 0.425921231508255
step: 70, loss: 0.40079644322395325
step: 80, loss: 0.2043275088071823
step: 90, loss: 0.2946349084377289
step: 100, loss: 0.5511829853057861
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.23382124304771423
step: 120, loss: 0.3387622833251953
step: 130, loss: 0.3292425274848938
step: 140, loss: 0.15907001495361328
step: 150, loss: 0.3264100253582001
step: 160, loss: 0.23375825583934784
step: 170, loss: 0.42022472620010376
step: 180, loss: 0.15770551562309265
step: 190, loss: 0.1522648185491562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20300751879699247, f1=0.2064864864864865,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3845862150192261
step: 10, loss: 0.3354078531265259
step: 20, loss: 0.8552823066711426
step: 30, loss: 0.3120412230491638
step: 40, loss: 0.597998857498169
```

```
step: 50, loss: 0.2898830771446228
step: 60, loss: 0.4958529472351074
step: 70, loss: 0.3080691397190094
step: 80, loss: 0.1628357470035553
step: 90, loss: 0.29936426877975464
step: 100, loss: 0.21618007123470306
step: 110, loss: 0.3646174967288971
step: 120, loss: 0.22380319237709045
step: 130, loss: 0.4654853940010071
step: 140, loss: 0.30026865005493164
step: 150, loss: 0.29605764150619507
step: 160, loss: 0.2770642340183258
step: 170, loss: 0.2378309816122055
step: 180, loss: 0.17886626720428467
step: 190, loss: 0.22754761576652527
/usr/local/lib/python3.7/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=0.20075963103635378, f1=0.20479302832244012,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3553221523761749
step: 10, loss: 0.35711395740509033
step: 20, loss: 0.4403833746910095
step: 30, loss: 0.2972676455974579
step: 40, loss: 0.08800748735666275
step: 50, loss: 0.3855598568916321
step: 60, loss: 0.16502726078033447
step: 70, loss: 0.3767458498477936
step: 80, loss: 0.2941255271434784
step: 90, loss: 0.36990389227867126
step: 100, loss: 0.5069512724876404
step: 110, loss: 0.6747727990150452
step: 120, loss: 0.3658531606197357
step: 130, loss: 0.16032367944717407
step: 140, loss: 0.360663503408432
step: 150, loss: 0.35795846581459045
step: 160, loss: 0.587875247001648
step: 170, loss: 0.4055929183959961
step: 180, loss: 0.3558028042316437
step: 190, loss: 0.16739997267723083
/usr/local/lib/python3.7/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=0.20075963103635378, f1=0.20479302832244012,
best_f1=0.2064864864864865
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2327776998281479
step: 10, loss: 0.22684720158576965
step: 20, loss: 0.27414044737815857
step: 30, loss: 0.23338714241981506
step: 40, loss: 0.5205308198928833
step: 50, loss: 0.22434799373149872
step: 60, loss: 0.346368670463562
step: 70, loss: 0.28561851382255554
step: 80, loss: 0.20282986760139465
step: 90, loss: 0.1525849997997284
step: 100, loss: 0.38052818179130554
step: 110, loss: 0.4306599199771881
step: 120, loss: 0.22437737882137299
step: 130, loss: 0.412801593542099
step: 140, loss: 0.38397496938705444
step: 150, loss: 0.2131388783454895
step: 160, loss: 0.29655134677886963
step: 170, loss: 0.42463597655296326
step: 180, loss: 0.367583692073822
step: 190, loss: 0.15065069496631622
/usr/local/lib/python3.7/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=0.20075963103635378, f1=0.20479302832244012,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36166971921920776
step: 10, loss: 0.3787503242492676
step: 20, loss: 0.14026997983455658
step: 30, loss: 0.12370806932449341
step: 40, loss: 0.3009544610977173
step: 50, loss: 0.48438113927841187
step: 60, loss: 0.2350471317768097
step: 70, loss: 0.36130934953689575
step: 80, loss: 0.33996301889419556
step: 90, loss: 0.2854309380054474
step: 100, loss: 0.4056668281555176
step: 110, loss: 0.3696691393852234
step: 120, loss: 0.23514911532402039
step: 130, loss: 0.500726044178009
step: 140, loss: 0.34937071800231934
step: 150, loss: 0.2856294810771942
step: 160, loss: 0.1570524275302887
```

```
step: 170, loss: 0.3497491180896759
step: 180, loss: 0.23980174958705902
step: 190, loss: 0.28892478346824646
/usr/local/lib/python3.7/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=0.20075963103635378, f1=0.20479302832244012,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2845560312271118
step: 10, loss: 0.23549139499664307
step: 20, loss: 0.28851407766342163
step: 30, loss: 0.450594961643219
step: 40, loss: 0.2840299904346466
step: 50, loss: 0.3103931248188019
step: 60, loss: 0.3879595696926117
step: 70, loss: 0.27534887194633484
step: 80, loss: 0.2821000814437866
step: 90, loss: 0.2196420133113861
step: 100, loss: 0.4429841637611389
step: 110, loss: 0.2196992188692093
step: 120, loss: 0.44818875193595886
step: 130, loss: 0.5454692244529724
step: 140, loss: 0.1791856735944748
step: 150, loss: 0.3620415925979614
step: 160, loss: 0.3686161935329437
step: 170, loss: 0.36277180910110474
step: 180, loss: 0.18083754181861877
step: 190, loss: 0.3105088174343109
/usr/local/lib/python3.7/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=0.20150862068965517, f1=0.20521172638436483,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3075127899646759
step: 10, loss: 0.2825274169445038
step: 20, loss: 0.2870886027812958
step: 30, loss: 0.20121169090270996
step: 40, loss: 0.3016749620437622
step: 50, loss: 0.09456562995910645
step: 60, loss: 0.1511499285697937
step: 70, loss: 0.1459200382232666
step: 80, loss: 0.21906687319278717
```

```
step: 90, loss: 0.23722892999649048
step: 100, loss: 0.5479375720024109
step: 110, loss: 0.47999781370162964
step: 120, loss: 0.35218775272369385
step: 130, loss: 0.29418161511421204
step: 140, loss: 0.22213388979434967
step: 150, loss: 0.27645835280418396
step: 160, loss: 0.36688458919525146
step: 170, loss: 0.3067454695701599
step: 180, loss: 0.21025845408439636
step: 190, loss: 0.30174532532691956
/usr/local/lib/python3.7/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=0.20075963103635378, f1=0.20479302832244012,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22845081984996796
step: 10, loss: 0.35264381766319275
step: 20, loss: 0.22236691415309906
step: 30, loss: 0.3485003113746643
step: 40, loss: 0.15328550338745117
step: 50, loss: 0.3028590679168701
step: 60, loss: 0.4332078695297241
step: 70, loss: 0.22737297415733337
step: 80, loss: 0.3483816683292389
step: 90, loss: 0.23536930978298187
step: 100, loss: 0.2918989360332489
step: 110, loss: 0.3780948221683502
step: 120, loss: 0.36133047938346863
step: 130, loss: 0.29620128870010376
step: 140, loss: 0.3616560399532318
step: 150, loss: 0.3020448088645935
step: 160, loss: 0.17511965334415436
step: 170, loss: 0.22685933113098145
step: 180, loss: 0.3078194856643677
step: 190, loss: 0.2666730582714081
/usr/local/lib/python3.7/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=0.20075963103635378, f1=0.20479302832244012,
best_f1=0.2064864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22213555872440338
```

```
step: 10, loss: 0.09364079684019089
step: 20, loss: 0.3040192127227783
step: 30, loss: 0.16598914563655853
step: 40, loss: 0.3116125166416168
step: 50, loss: 0.3698103427886963
step: 60, loss: 0.36319854855537415
step: 70, loss: 0.15750464797019958
step: 80, loss: 0.3542887270450592
step: 90, loss: 0.7478327751159668
step: 100, loss: 0.3502029478549957
step: 110, loss: 0.3416993319988251
step: 120, loss: 0.5548515319824219
step: 130, loss: 0.2976970970630646
step: 140, loss: 0.28732478618621826
step: 150, loss: 0.24252858757972717
step: 160, loss: 0.21458671987056732
step: 170, loss: 0.5428652167320251
step: 180, loss: 0.37334147095680237
step: 190, loss: 0.23440293967723846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20324324324324325, f1=0.2044589450788472,
best_f1=0.2044589450788472
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2712841331958771
step: 10, loss: 0.1633705049753189
step: 20, loss: 0.3065645098686218
step: 30, loss: 0.36353132128715515
step: 40, loss: 0.29450497031211853
step: 50, loss: 0.09579361230134964
step: 60, loss: 0.2997792065143585
step: 70, loss: 0.289974570274353
step: 80, loss: 0.36024031043052673
step: 90, loss: 0.13856078684329987
step: 100, loss: 0.21194837987422943
step: 110, loss: 0.22064460813999176
step: 120, loss: 0.3010956645011902
step: 130, loss: 0.4066815674304962
step: 140, loss: 0.3168693780899048
step: 150, loss: 0.2870579659938812
step: 160, loss: 0.12198273837566376
```

```
step: 170, loss: 0.2032766193151474
step: 180, loss: 0.15889255702495575
step: 190, loss: 0.10645988583564758
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6944444444444444, f1=0.6857142857142857,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30892279744148254
step: 10, loss: 0.05196506530046463
step: 20, loss: 0.09629163891077042
step: 30, loss: 0.21372519433498383
step: 40, loss: 0.053853075951337814
step: 50, loss: 0.1732044517993927
step: 60, loss: 0.1317947506904602
step: 70, loss: 0.18791957199573517
step: 80, loss: 0.37713390588760376
step: 90, loss: 0.22503511607646942
step: 100, loss: 0.16089005768299103
step: 110, loss: 0.20129768550395966
step: 120, loss: 0.1803346425294876
step: 130, loss: 0.2928487956523895
step: 140, loss: 0.1778784543275833
step: 150, loss: 0.038753315806388855
step: 160, loss: 0.22494558990001678
step: 170, loss: 0.3072911202907562
step: 180, loss: 0.06617988646030426
step: 190, loss: 0.054136086255311966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7846153846153846, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12975859642028809
```

```
step: 10, loss: 0.09983011335134506
step: 20, loss: 0.12292838096618652
step: 30, loss: 0.18687841296195984
step: 40, loss: 0.05410280451178551
step: 50, loss: 0.23007921874523163
step: 60, loss: 0.1760353147983551
step: 70, loss: 0.07007288187742233
step: 80, loss: 0.2343740165233612
step: 90, loss: 0.14564888179302216
step: 100, loss: 0.17151129245758057
step: 110, loss: 0.3561105728149414
step: 120, loss: 0.05932978168129921
step: 130, loss: 0.03791479393839836
step: 140, loss: 0.02969687432050705
step: 150, loss: 0.18230359256267548
step: 160, loss: 0.05650990828871727
step: 170, loss: 0.24834421277046204
step: 180, loss: 0.0851273462176323
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.14333893358707428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8091168091168092, f1=0.7988668555240793,
best_f1=0.7988668555240793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18312394618988037
step: 10, loss: 0.12494039535522461
step: 20, loss: 0.18876305222511292
step: 30, loss: 0.34995216131210327
step: 40, loss: 0.016110308468341827
step: 50, loss: 0.011762960813939571
step: 60, loss: 0.08791971951723099
step: 70, loss: 0.04945032671093941
step: 80, loss: 0.0923505648970604
step: 90, loss: 0.1999988555908203
step: 100, loss: 0.019709551706910133
step: 110, loss: 0.0980406329035759
step: 120, loss: 0.018829768523573875
step: 130, loss: 0.1310051828622818
step: 140, loss: 0.16606763005256653
step: 150, loss: 0.1114170178771019
```

```
step: 160, loss: 0.09349571913480759
step: 170, loss: 0.09282047301530838
step: 180, loss: 0.021488824859261513
step: 190, loss: 0.262574076655252686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8306878306878307, f1=0.8186813186813187,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03440660983324051
step: 10, loss: 0.033440571278333664
step: 20, loss: 0.08378192782402039
step: 30, loss: 0.03354981914162636
step: 40, loss: 0.03360140323638916
step: 50, loss: 0.07168523222208023
step: 60, loss: 0.040794357657432556
step: 70, loss: 0.02266114018857479
step: 80, loss: 0.032322295010089874
step: 90, loss: 0.047021958976984024
step: 100, loss: 0.05987461656332016
step: 110, loss: 0.027527179569005966
step: 120, loss: 0.10713040083646774
step: 130, loss: 0.015154790133237839
step: 140, loss: 0.05603121221065521
step: 150, loss: 0.09430979937314987
step: 160, loss: 0.2366810292005539
step: 170, loss: 0.026799581944942474
step: 180, loss: 0.07216013222932816
step: 190, loss: 0.27802538871765137
/usr/local/lib/python3.7/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=0.8300835654596099, f1=0.8135593220338984,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13200069963932037
step: 10, loss: 0.09149452298879623
step: 20, loss: 0.14304229617118835
step: 30, loss: 0.03899357095360756
```

```
step: 40, loss: 0.026376305148005486
step: 50, loss: 0.017942512407898903
step: 60, loss: 0.017595959827303886
step: 70, loss: 0.14566445350646973
step: 80, loss: 0.16907905042171478
step: 90, loss: 0.2378605753183365
step: 100, loss: 0.1435912847518921
step: 110, loss: 0.10343357175588608
step: 120, loss: 0.03712878003716469
step: 130, loss: 0.049884337931871414
step: 140, loss: 0.20833832025527954
step: 150, loss: 0.018438108265399933
step: 160, loss: 0.011757693253457546
step: 170, loss: 0.2152996063232422
step: 180, loss: 0.01323855109512806
step: 190, loss: 0.22638389468193054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.8342541436464089, f1=0.8133704735376045,
best_f1=0.8133704735376045
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 [00:09, 212.61it/s]
load_f1 = 0.8283378746594007
real_f1 = 0.8263157894736842
733it [00:00, 3223.55it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 177.54it/s]
```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

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

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

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

```
Selected optimization level 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: 0.4939292371273041
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.42542892694473267
step: 20, loss: 0.28638148307800293
step: 30, loss: 0.4758548140525818
step: 40, loss: 0.48360151052474976
step: 50, loss: 0.30877989530563354
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.5663739442825317
step: 70, loss: 0.32669132947921753
step: 80, loss: 0.28699061274528503
step: 90, loss: 0.2886590361595154
```

```
step: 100, loss: 0.16096583008766174
step: 110, loss: 0.44427433609962463
step: 120, loss: 0.3442075252532959
step: 130, loss: 0.2774694263935089
step: 140, loss: 0.3847621977329254
step: 150, loss: 0.3300885558128357
step: 160, loss: 0.42018958926200867
step: 170, loss: 0.29409730434417725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3026315789473684, f1=0.25407166123778496,
best_f1=0.25407166123778496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26725029945373535
step: 10, loss: 0.4337259531021118
step: 20, loss: 0.3150181174278259
step: 30, loss: 0.2630099952220917
step: 40, loss: 0.04888302460312843
step: 50, loss: 0.2642747461795807
step: 60, loss: 0.1254301369190216
step: 70, loss: 0.26158085465431213
step: 80, loss: 0.1489696204662323
step: 90, loss: 0.06273036450147629
step: 100, loss: 0.23030737042427063
step: 110, loss: 0.09441566467285156
step: 120, loss: 0.03462040796875954
step: 130, loss: 0.24203939735889435
step: 140, loss: 0.23354361951351166
step: 150, loss: 0.08475538343191147
step: 160, loss: 0.09899018704891205
step: 170, loss: 0.09076298028230667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7901234567901234, f1=0.7925407925407925,
best_f1=0.7925407925407925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20571765303611755
step: 10, loss: 0.08943554013967514
step: 20, loss: 0.08360151946544647
step: 30, loss: 0.049034640192985535
step: 40, loss: 0.16695217788219452
step: 50, loss: 0.27053409814834595
step: 60, loss: 0.03362613916397095
step: 70, loss: 0.07315661013126373
step: 80, loss: 0.2459302544593811
step: 90, loss: 0.233729287981987
step: 100, loss: 0.04248550906777382
step: 110, loss: 0.023406824097037315
step: 120, loss: 0.03678475692868233
step: 130, loss: 0.11018290370702744
step: 140, loss: 0.17078204452991486
step: 150, loss: 0.0844138115644455
step: 160, loss: 0.07032511383295059
step: 170, loss: 0.10521067678928375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8030690537084398, f1=0.8243902439024391,
best_f1=0.8243902439024391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09970706701278687
step: 10, loss: 0.07141975313425064
step: 20, loss: 0.007908729836344719
step: 30, loss: 0.046493466943502426
step: 40, loss: 0.034606728702783585
step: 50, loss: 0.07011397182941437
step: 60, loss: 0.05496850982308388
step: 70, loss: 0.002487157704308629
step: 80, loss: 0.45588621497154236
step: 90, loss: 0.3895372748374939
step: 100, loss: 0.1965375393629074
step: 110, loss: 0.050692118704319
step: 120, loss: 0.37608957290649414
step: 130, loss: 0.08067416399717331
step: 140, loss: 0.10500365495681763
step: 150, loss: 0.170075923204422
```

```
step: 160, loss: 0.002712022978812456
step: 170, loss: 0.043537888675928116
/usr/local/lib/python3.7/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=0.786096256684492, f1=0.8081841432225064,
best_f1=0.8243902439024391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021352458745241165
step: 10, loss: 0.0403513722121171555
step: 20, loss: 0.0442086756229400063
step: 30, loss: 0.0025936360470950603
step: 40, loss: 0.018442152068018913
step: 50, loss: 0.002598607214167714
step: 60, loss: 0.04051833599805832
step: 70, loss: 0.01997048407793045
step: 80, loss: 0.0005108332261443138
step: 90, loss: 0.007882134057581425
step: 100, loss: 0.029117928817868233
step: 110, loss: 0.06127266585826874
step: 120, loss: 0.18263255059719086
step: 130, loss: 0.003560310462489724
step: 140, loss: 0.036183107644319534
step: 150, loss: 0.05054215341806412
step: 160, loss: 0.08244699984788895
step: 170, loss: 0.016355128958821297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8211586901763224, f1=0.8465346534653465,
best_f1=0.8465346534653465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08274364471435547
step: 10, loss: 0.018342822790145874
step: 20, loss: 0.027058707550168037
step: 30, loss: 0.0016379448352381587
step: 40, loss: 0.0032630981877446175
step: 50, loss: 0.00589147862046957
step: 60, loss: 0.02645624615252018
step: 70, loss: 0.012810979038476944
```

```
step: 80, loss: 0.00798901729285717
step: 90, loss: 0.03322276100516319
step: 100, loss: 0.04058723524212837
step: 110, loss: 0.05937829986214638
step: 120, loss: 0.005917403381317854
step: 130, loss: 0.07020986080169678
step: 140, loss: 0.06079281121492386
step: 150, loss: 0.0056083593517541885
step: 160, loss: 0.0016146480338647962
step: 170, loss: 0.006966139655560255
/usr/local/lib/python3.7/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=0.8190954773869347, f1=0.8238095238095239,
best_f1=0.8465346534653465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01004233118146658
step: 10, loss: 0.00811281893402338
step: 20, loss: 0.002816049847751856
step: 30, loss: 0.0013006163062527776
step: 40, loss: 0.0005831732414662838
step: 50, loss: 0.0006704297848045826
step: 60, loss: 0.007485878653824329
step: 70, loss: 0.002743223449215293
step: 80, loss: 0.009062951430678368
step: 90, loss: 0.03988904505968094
step: 100, loss: 0.0010596414795145392
step: 110, loss: 0.1508331447839737
step: 120, loss: 0.0045747836120426655
step: 130, loss: 0.014069583266973495
step: 140, loss: 0.02821798250079155
step: 150, loss: 0.2061462104320526
step: 160, loss: 0.011224297806620598
step: 170, loss: 0.048373762518167496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.824742268041237, f1=0.8188585607940446,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0087279649451375
step: 10, loss: 0.02251950092613697
step: 20, loss: 0.008865993469953537
step: 30, loss: 0.00045284017687663436
step: 40, loss: 0.0450027696788311
step: 50, loss: 0.0022027783561497927
step: 60, loss: 0.013920195400714874
step: 70, loss: 0.04790298640727997
step: 80, loss: 0.0023757461458444595
step: 90, loss: 0.01337698008865118
step: 100, loss: 0.0009260991355404258
step: 110, loss: 0.05048273503780365
step: 120, loss: 0.009141549468040466
step: 130, loss: 0.001594026107341051
step: 140, loss: 0.015052353031933308
step: 150, loss: 0.005543706472963095
step: 160, loss: 0.0006272721802815795
step: 170, loss: 0.1962910294532776
/usr/local/lib/python3.7/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=0.813131313131313, f1=0.8321167883211679,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0305425263941288
step: 10, loss: 0.0013553687604144216
step: 20, loss: 0.010641190223395824
step: 30, loss: 0.0003865724429488182
step: 40, loss: 0.05294543132185936
step: 50, loss: 0.0007312792586162686
step: 60, loss: 0.02360599860548973
step: 70, loss: 0.13614612817764282
step: 80, loss: 0.0009084048797376454
step: 90, loss: 0.029408618807792664
step: 100, loss: 0.00872716773301363
step: 110, loss: 0.15335150063037872
step: 120, loss: 0.02573719434440136
step: 130, loss: 0.0002920961705967784
step: 140, loss: 0.00021878523693885654
step: 150, loss: 0.19210252165794373
step: 160, loss: 0.006921758875250816
step: 170, loss: 0.008536563254892826
/usr/local/lib/python3.7/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=0.8123393316195373, f1=0.8275862068965517,
```

```
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007185556925833225
step: 10, loss: 0.004367881454527378
step: 20, loss: 0.05443023890256882
step: 30, loss: 0.047733381390571594
step: 40, loss: 0.002619100036099553
step: 50, loss: 0.034867290407419205
step: 60, loss: 0.009628722444176674
step: 70, loss: 0.00037054589483886957
step: 80, loss: 0.016410276293754578
step: 90, loss: 0.0011329582193866372
step: 100, loss: 0.002588655101135373
step: 110, loss: 0.010632598772644997
step: 120, loss: 0.0037170113064348698
step: 130, loss: 0.009162209928035736
step: 140, loss: 0.0038435417227447033
step: 150, loss: 0.00023831047292333096
step: 160, loss: 0.00032782458583824337
step: 170, loss: 0.023272467777132988
/usr/local/lib/python3.7/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=0.8198433420365536, f1=0.8261964735516373,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19484475255012512
step: 10, loss: 0.0002795292530208826
step: 20, loss: 0.00011465014540590346
step: 30, loss: 0.005855193827301264
step: 40, loss: 0.0012683210661634803
step: 50, loss: 0.0003667572746053338
step: 60, loss: 0.001335035776719451
step: 70, loss: 0.00032834135345183313
step: 80, loss: 0.022421542555093765
step: 90, loss: 0.0014684797497466207
step: 100, loss: 0.021890491247177124
step: 110, loss: 0.0002312627766514197
step: 120, loss: 0.00010425061918795109
step: 130, loss: 0.011945068836212158
step: 140, loss: 0.0019795678090304136
step: 150, loss: 0.0002798197092488408
step: 160, loss: 0.03761770576238632
step: 170, loss: 0.0004014185396954417
```

```
/usr/local/lib/python3.7/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=0.8241469816272966, f1=0.8303797468354431,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005878923693671823
step: 10, loss: 0.0001370625977870077
step: 20, loss: 0.00015149146202020347
step: 30, loss: 0.04784752056002617
step: 40, loss: 0.0005878000520169735
step: 50, loss: 0.00044027098920196295
step: 60, loss: 0.00027080695144832134
step: 70, loss: 6.870181096019223e-05
step: 80, loss: 0.00012310186866670847
step: 90, loss: 0.0009150532423518598
step: 100, loss: 0.00039905705489218235
step: 110, loss: 0.045658182352781296
step: 120, loss: 0.006296972744166851
step: 130, loss: 0.0014678790466859937
step: 140, loss: 9.926394704962149e-05
step: 150, loss: 0.0003805134037975222
step: 160, loss: 0.003566573141142726
step: 170, loss: 0.010619674809277058
/usr/local/lib/python3.7/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=0.817258883248731, f1=0.8341463414634146,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012621954083442688
step: 10, loss: 0.007537521421909332
step: 20, loss: 0.00010653387289494276
step: 30, loss: 0.045112669467926025
step: 40, loss: 0.00012472676462493837
step: 50, loss: 0.008640528656542301
step: 60, loss: 0.00035352600389160216
step: 70, loss: 0.02166062779724598
step: 80, loss: 0.00020835423492826521
step: 90, loss: 0.00025843738694675267
step: 100, loss: 0.013332301750779152
step: 110, loss: 0.0003610215790104121
step: 120, loss: 0.007317183073610067
step: 130, loss: 4.700034696725197e-05
```

```
step: 140, loss: 0.020329132676124573
step: 150, loss: 8.147647167788818e-05
step: 160, loss: 8.649436495034024e-05
step: 170, loss: 0.0002521765709389001
/usr/local/lib/python3.7/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=0.8177083333333334, f1=0.8235294117647058,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012166839587735012
step: 10, loss: 4.6288769226521254e-05
step: 20, loss: 0.016375703737139702
step: 30, loss: 0.006191631779074669
step: 40, loss: 0.0002675464784260839
step: 50, loss: 0.0027748493012040854
step: 60, loss: 0.0004213734355289489
step: 70, loss: 0.00020219072757754475
step: 80, loss: 4.075330798514187e-05
step: 90, loss: 9.528710506856441e-05
step: 100, loss: 6.934799603186548e-05
step: 110, loss: 0.01018914207816124
step: 120, loss: 7.871657726354897e-05
step: 130, loss: 0.036576781421899796
step: 140, loss: 0.0012495772680267692
step: 150, loss: 0.021421702578663826
step: 160, loss: 0.0008744837832637131
step: 170, loss: 0.03440452739596367
/usr/local/lib/python3.7/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=0.8126649076517151, f1=0.8229426433915212,
best_f1=0.8188585607940446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.682512008002959e-05
step: 10, loss: 0.0003431156510487199
step: 20, loss: 7.364212069660425e-05
step: 30, loss: 0.0003164085792377591
step: 40, loss: 0.0280440766364336
step: 50, loss: 6.124427454778925e-05
step: 60, loss: 5.784295717603527e-05
step: 70, loss: 0.012110949493944645
step: 80, loss: 9.776778460945934e-05
step: 90, loss: 8.79926883499138e-05
```

```
step: 100, loss: 0.06386376172304153
step: 110, loss: 0.00011416708002798259
step: 120, loss: 9.063955076271668e-05
step: 130, loss: 0.0011532228672876954
step: 140, loss: 0.0072428984567523
step: 150, loss: 0.0013529288116842508
step: 160, loss: 5.430772580439225e-05
step: 170, loss: 0.0004759363073389977
/usr/local/lib/python3.7/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=0.8157894736842105, f1=0.8238213399503721,
best_f1=0.8188585607940446
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, 234.66it/s]
load_f1 = 0.824742268041237
real_f1 = 0.8277634961439588
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 146.47it/s]
```

2.3 BASELINE DIRTY

2.3.1 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 roberta \
  --fp16 \
  --save_model

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

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

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

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
```

```
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5949551463127136
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4445851147174835
step: 20, loss: 0.49952316284179695
step: 30, loss: 0.34585970640182495
step: 40, loss: 0.3370807468891144
step: 50, loss: 0.5389039516448975
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.46804410219192505
step: 70, loss: 0.42100629210472107
step: 80, loss: 0.6263421177864075
step: 90, loss: 0.4485740065574646
step: 100, loss: 0.46742475032806396
step: 110, loss: 0.5210608243942261
step: 120, loss: 0.22983747720718384
step: 130, loss: 0.13574349880218506
step: 140, loss: 0.2034178227186203
step: 150, loss: 0.17912811040878296
step: 160, loss: 0.1725236028432846
step: 170, loss: 0.16531455516815186
step: 180, loss: 0.0845450833439827
step: 190, loss: 0.10258016735315323
step: 200, loss: 0.05345602333545685
step: 210, loss: 0.1553088128566742
step: 220, loss: 0.1300457864999771
step: 230, loss: 0.011428000405430794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9311111111111112, f1=0.9548532731376976,
best_f1=0.9548532731376976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01175769604742527
step: 10, loss: 0.13674257695674896
step: 20, loss: 0.017206551507115364
step: 30, loss: 0.020385710522532463
step: 40, loss: 0.058134905993938446
step: 50, loss: 0.06849272549152374
step: 60, loss: 0.04906633123755455
step: 70, loss: 0.011626888066530228
step: 80, loss: 0.003642293158918619
step: 90, loss: 0.011830279603600502
step: 100, loss: 0.1569506973028183
step: 110, loss: 0.2341635674238205
step: 120, loss: 0.008943265303969383
step: 130, loss: 0.04263210669159889
step: 140, loss: 0.016223374754190445
step: 150, loss: 0.13633324205875397
step: 160, loss: 0.05230433866381645
step: 170, loss: 0.011070254258811474
step: 180, loss: 0.01174250990152359
step: 190, loss: 0.015992525964975357
step: 200, loss: 0.02019045129418373
step: 210, loss: 0.0928693413734436
step: 220, loss: 0.004314372781664133
step: 230, loss: 0.0015203990042209625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9740698985343857, f1=0.9577142857142856,
best_f1=0.9577142857142856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05261286348104477
step: 10, loss: 0.004236321896314621
step: 20, loss: 0.03282987326383591
step: 30, loss: 0.01129295863211155
step: 40, loss: 0.047557901591062546
step: 50, loss: 0.0053245085291564465
```

```
step: 60, loss: 0.03990750387310982
step: 70, loss: 0.0017068138113245368
step: 80, loss: 0.12092866748571396
step: 90, loss: 0.00763006042689085
step: 100, loss: 0.0007379449089057744
step: 110, loss: 0.01649031788110733
step: 120, loss: 0.0010758851421996951
step: 130, loss: 0.031684406101703644
step: 140, loss: 0.01666220650076866
step: 150, loss: 0.0457390658557415
step: 160, loss: 0.0014931178884580731
step: 170, loss: 0.02858547866344452
step: 180, loss: 0.0355987511575222
step: 190, loss: 0.06020976975560188
step: 200, loss: 0.0061628553085029125
step: 210, loss: 0.0030757079366594553
step: 220, loss: 0.1106298640370369
step: 230, loss: 0.003276634495705366
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.975609756097561, f1=0.9654403567447045,
best_f1=0.9654403567447045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0385926179587841
step: 10, loss: 0.0024610874243080616
step: 20, loss: 0.01969761960208416
step: 30, loss: 0.0020753098651766777
step: 40, loss: 0.19035986065864563
step: 50, loss: 0.09224364161491394
step: 60, loss: 0.013662080280482769
step: 70, loss: 0.023513641208410263
step: 80, loss: 0.010053505189716816
step: 90, loss: 0.015871092677116394
step: 100, loss: 0.009537145495414734
step: 110, loss: 0.0009280507219955325
step: 120, loss: 0.01228982675820589
step: 130, loss: 0.005644838325679302
step: 140, loss: 0.004753750748932362
step: 150, loss: 0.009042215533554554
step: 160, loss: 0.06493326276540756
step: 170, loss: 0.006633851211518049
```

```
step: 180, loss: 0.002297263126820326
step: 190, loss: 0.0013564751716330647
step: 200, loss: 0.09946125745773315
step: 210, loss: 0.01045567449182272
step: 220, loss: 0.0010195884387940168
step: 230, loss: 0.0015839484985917807
/usr/local/lib/python3.7/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=0.9755011135857461, f1=0.9707865168539327,
best_f1=0.9654403567447045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006647862028330564
step: 10, loss: 0.002718631410971284
step: 20, loss: 0.00105145329143852
step: 30, loss: 0.0007121968665160239
step: 40, loss: 0.0017713578417897224
step: 50, loss: 0.0026752145495265722
step: 60, loss: 0.04822375252842903
step: 70, loss: 0.007584588602185249
step: 80, loss: 0.04526408389210701
step: 90, loss: 0.07770729809999466
step: 100, loss: 0.00812316033989191
step: 110, loss: 0.007713480852544308
step: 120, loss: 0.020118195563554764
step: 130, loss: 0.00745819229632616
step: 140, loss: 0.014872951433062553
step: 150, loss: 0.014247854240238667
step: 160, loss: 0.0006454398389905691
step: 170, loss: 0.019541623070836067
step: 180, loss: 0.004133300390094519
step: 190, loss: 0.00228700484149158
step: 200, loss: 0.01747230254113674
step: 210, loss: 0.023705536499619484
step: 220, loss: 0.002910703420639038
step: 230, loss: 0.009873435832560062
/usr/local/lib/python3.7/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=0.9754464285714286, f1=0.9648924122310306,
best_f1=0.9654403567447045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006180083379149437
step: 10, loss: 0.0008731192210689187
```

```
step: 20, loss: 0.0029414750169962645
step: 30, loss: 0.0013187731383368373
step: 40, loss: 0.0002574855461716652
step: 50, loss: 0.001515998737886548
step: 60, loss: 0.0018619367619976401
step: 70, loss: 0.08396937698125839
step: 80, loss: 0.0007442728383466601
step: 90, loss: 0.004588214214891195
step: 100, loss: 0.004618730396032333
step: 110, loss: 0.008121935650706291
step: 120, loss: 0.002443678444251418
step: 130, loss: 0.0005461851833388209
step: 140, loss: 0.004589785821735859
step: 150, loss: 0.00018507425556890666
step: 160, loss: 0.00022678301320411265
step: 170, loss: 0.00014438465586863458
step: 180, loss: 0.00046605963143520057
step: 190, loss: 0.0003027792845387012
step: 200, loss: 0.0006039371946826577
step: 210, loss: 0.0003816426615230739
step: 220, loss: 0.019930768758058548
step: 230, loss: 0.0003500028979033232
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9786276715410572, f1=0.9739524348810873,
best_f1=0.9739524348810873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005619055591523647
step: 10, loss: 0.00021410519548226148
step: 20, loss: 0.00015586658264510334
step: 30, loss: 0.00015968549996614456
step: 40, loss: 0.0005569742643274367
step: 50, loss: 0.0003348525206092745
step: 60, loss: 0.00025618108338676393
step: 70, loss: 0.0005995925166644156
step: 80, loss: 0.06307297199964523
step: 90, loss: 0.0007590571767650545
step: 100, loss: 0.00015663467638660222
step: 110, loss: 0.00019457894086372107
step: 120, loss: 0.0018812651978805661
step: 130, loss: 0.031586069613695145
```

```
step: 140, loss: 0.0002962685830425471
step: 150, loss: 0.02504115179181099
step: 160, loss: 0.009240575134754181
step: 170, loss: 0.02150075137615204
step: 180, loss: 0.0003104100760538131
step: 190, loss: 0.0005143416346982121
step: 200, loss: 0.0016614858759567142
step: 210, loss: 0.024948997423052788
step: 220, loss: 0.0020729240495711565
step: 230, loss: 0.007710180711001158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787234042553192, f1=0.971687429218573,
best_f1=0.971687429218573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013435387518256903
step: 10, loss: 0.0014262403128668666
step: 20, loss: 0.0002724316145759076
step: 30, loss: 0.004549143835902214
step: 40, loss: 0.0009480596636421978
step: 50, loss: 0.005400713998824358
step: 60, loss: 0.0003709607117343694
step: 70, loss: 0.00016997862257994711
step: 80, loss: 0.006661179009824991
step: 90, loss: 0.0007210171897895634
step: 100, loss: 0.00020719249732792377
step: 110, loss: 0.0027568484656512737
step: 120, loss: 9.68187814578414e-05
step: 130, loss: 0.00020213128300383687
step: 140, loss: 0.00012705016706660451
step: 150, loss: 0.16808825731277466
step: 160, loss: 0.0006783555145375431
step: 170, loss: 0.03025970049202442
step: 180, loss: 0.00047515929327346385
step: 190, loss: 0.024311767891049385
step: 200, loss: 0.002181839197874069
step: 210, loss: 0.0008697470766492188
step: 220, loss: 0.00033865327714011073
step: 230, loss: 0.000248458149144426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.980963045912654, f1=0.9796380090497738,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011918210657313466
step: 10, loss: 0.005308710969984531
step: 20, loss: 0.00013955739268567413
step: 30, loss: 0.0002566208131611347
step: 40, loss: 8.196067938115448e-05
step: 50, loss: 0.0011700839968398213
step: 60, loss: 0.00010457566531840712
step: 70, loss: 0.00038099681842140853
step: 80, loss: 6.38637357042171e-05
step: 90, loss: 0.07232566922903061
step: 100, loss: 7.917365292087197e-05
step: 110, loss: 8.848913421388716e-05
step: 120, loss: 0.0004388346860650927
step: 130, loss: 0.00017325991939287633
step: 140, loss: 0.0001715376420179382
step: 150, loss: 0.0002763614465948194
step: 160, loss: 0.00029020459624007344
step: 170, loss: 9.693354513728991e-05
step: 180, loss: 0.0018812627531588078
step: 190, loss: 0.00010115613258676603
step: 200, loss: 0.002644206862896681
step: 210, loss: 0.015168345533311367
step: 220, loss: 0.00025392998941242695
step: 230, loss: 0.00021448491315823048
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9774774774774775,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018696188635658473
step: 10, loss: 0.00046897598076611757
step: 20, loss: 0.0003143279755022377
step: 30, loss: 0.00018581448239274323
step: 40, loss: 0.0008561840513721108
step: 50, loss: 0.00022721040295436978
```

```
step: 60, loss: 0.00033930985955521464
step: 70, loss: 0.00044501142110675573
step: 80, loss: 0.00018725938571151346
step: 90, loss: 0.00011973567598033696
step: 100, loss: 0.00015181992785073817
step: 110, loss: 0.00018165666551794857
step: 120, loss: 0.00016193828196264803
step: 130, loss: 0.00024383632990065962
step: 140, loss: 0.00010896628373302519
step: 150, loss: 0.0004137771320529282
step: 160, loss: 6.900665175635368e-05
step: 170, loss: 0.0006812189240008593
step: 180, loss: 0.009682049043476582
step: 190, loss: 0.0004140371165703982
step: 200, loss: 0.0004604761488735676
step: 210, loss: 0.044050395488739014
step: 220, loss: 0.0004339482693467289
step: 230, loss: 0.00017205701442435384
/usr/local/lib/python3.7/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=0.9798206278026906, f1=0.9774266365688488,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011789921700255945
step: 10, loss: 0.0002627071226015687
step: 20, loss: 0.0001921315270010382
step: 30, loss: 0.015303111635148525
step: 40, loss: 2.792050872812979e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.0001206879605888389
step: 60, loss: 0.01759600266814232
step: 70, loss: 0.001328112673945725
step: 80, loss: 0.00012840265117119998
step: 90, loss: 0.0011280649341642857
step: 100, loss: 0.0002404141559964046
step: 110, loss: 0.0008946310263127089
step: 120, loss: 0.00022960569185670465
step: 130, loss: 0.00013401402975432575
step: 140, loss: 0.0005658186855725944
step: 150, loss: 0.00016513654554728419
step: 160, loss: 0.008465421386063099
step: 170, loss: 0.007373645901679993
step: 180, loss: 0.0008021368412300944
step: 190, loss: 0.0007367979851551354
step: 200, loss: 0.0014744398649781942
```

```
step: 210, loss: 0.0004517486959230155
step: 220, loss: 0.0009245658293366432
step: 230, loss: 0.00022719490516465157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9752252252252253, f1=0.9729119638826186,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004198843380436301
step: 10, loss: 0.00017627511988393962
step: 20, loss: 0.0007783471373841166
step: 30, loss: 0.0037863929755985737
step: 40, loss: 0.00037718238309025764
step: 50, loss: 0.0004047835827805102
step: 60, loss: 0.0009949945379048586
step: 70, loss: 0.00022958783665671945
step: 80, loss: 0.00011846996494568884
step: 90, loss: 0.0029904188122600317
step: 100, loss: 0.00024020115961320698
step: 110, loss: 0.0004300414293538779
step: 120, loss: 0.015511000528931618
step: 130, loss: 0.00028153625316917896
step: 140, loss: 0.00013576625497080386
step: 150, loss: 0.00012094457633793354
step: 160, loss: 0.000583620450925082
step: 170, loss: 0.0001298045681323856
step: 180, loss: 8.182731835404411e-05
step: 190, loss: 0.00012609500845428556
step: 200, loss: 0.0001240513811353594
step: 210, loss: 0.0002819678047671914
step: 220, loss: 0.0001071419901563786
step: 230, loss: 0.00014988277689553797
/usr/local/lib/python3.7/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=0.9798206278026906, f1=0.9763779527559054,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0046159126795828340
step: 10, loss: 0.00021935779659543186
step: 20, loss: 0.0002631754905451089
step: 30, loss: 0.0034563937224447727
step: 40, loss: 0.0005153220845386386
```

```
step: 50, loss: 0.0058837165124714375
step: 60, loss: 0.0007329025538638234
step: 70, loss: 0.00011949784675380215
step: 80, loss: 0.0009654498426243663
step: 90, loss: 0.00014746269152965397
step: 100, loss: 0.0002173012326238677
step: 110, loss: 0.00030381156830117106
step: 120, loss: 0.0003344025753904134
step: 130, loss: 0.0017983228899538517
step: 140, loss: 0.0006653160089626908
step: 150, loss: 9.481262532062829e-05
step: 160, loss: 0.0009642923250794411
step: 170, loss: 0.00036332852323539555
step: 180, loss: 0.0015659499913454056
step: 190, loss: 0.0003126698429696262
step: 200, loss: 5.364018943510018e-05
step: 210, loss: 0.00015951723617035896
step: 220, loss: 0.00017677500727586448
step: 230, loss: 0.00014402712986338884
/usr/local/lib/python3.7/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=0.9784824462061155, f1=0.9726651480637813,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011133213411085308
step: 10, loss: 6.260696682147682e-05
step: 20, loss: 0.00011255183926550671
step: 30, loss: 0.00031436531571671367
step: 40, loss: 0.0011860198574140668
step: 50, loss: 8.804374374449253e-05
step: 60, loss: 0.00013326921907719225
step: 70, loss: 0.0003950579557567835
step: 80, loss: 0.00011865151464007795
step: 90, loss: 0.00023189789499156177
step: 100, loss: 0.00011629102664301172
step: 110, loss: 0.00012172475544502959
step: 120, loss: 2.7598352971835993e-05
step: 130, loss: 0.00019236074876971543
step: 140, loss: 0.001692880759947002
step: 150, loss: 5.5795164371374995e-05
step: 160, loss: 0.00017703867342788726
step: 170, loss: 0.0001545740378787741
step: 180, loss: 6.9759989855811e-05
step: 190, loss: 4.8833684559213e-05
step: 200, loss: 0.00030443741707131267
```

```
step: 210, loss: 8.844218245940283e-05
step: 220, loss: 8.020576206035912e-05
step: 230, loss: 9.608100663172081e-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=0.9784824462061155, f1=0.9739524348810873,
best_f1=0.9796380090497738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000479000445920974
step: 10, loss: 8.465941937174648e-05
step: 20, loss: 0.00012650230200961232
step: 30, loss: 6.000822031637654e-05
step: 40, loss: 5.037850496592e-05
step: 50, loss: 4.81486385979224e-05
step: 60, loss: 4.79033260489814e-05
step: 70, loss: 0.001325066783465445
step: 80, loss: 0.00020651405793614686
step: 90, loss: 6.116661825217307e-05
step: 100, loss: 5.306801540427841e-05
step: 110, loss: 9.590679837856442e-05
step: 120, loss: 0.029949624091386795
step: 130, loss: 8.126220927806571e-05
step: 140, loss: 0.00015056112897582352
step: 150, loss: 0.00016678454994689673
step: 160, loss: 0.00014612250379286706
step: 170, loss: 4.275261744624004e-05
step: 180, loss: 7.681687566218898e-05
step: 190, loss: 4.4445874664234e-05
step: 200, loss: 0.00019884658104274422
step: 210, loss: 0.0009056419949047267
step: 220, loss: 6.809696787968278e-05
step: 230, loss: 0.00012068064097547904
/usr/local/lib/python3.7/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=0.9784824462061155, f1=0.9739524348810873,
best_f1=0.9796380090497738
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, 155.05it/s]
load_f1 = 0.980963045912654
real_f1 = 0.9798206278026906
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:29, 146.77it/s]
```

2.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6307740211486816
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.4194072484970093
step: 20, loss: 0.25823718309402466
step: 30, loss: 0.4127151072025299
```

```
step: 40, loss: 0.39976581931114197
step: 50, loss: 0.5458402633666992
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 60, loss: 0.35214048624038696
step: 70, loss: 0.3905305862426758
step: 80, loss: 0.20671029388904572
step: 90, loss: 0.14990614354610443
step: 100, loss: 0.3280629515647888
step: 110, loss: 0.1753973513841629
step: 120, loss: 0.1638607680797577
step: 130, loss: 0.19881604611873627
step: 140, loss: 0.24412134289741516
step: 150, loss: 0.02106526307761669
step: 160, loss: 0.24815452098846436
step: 170, loss: 0.10872111469507217
step: 180, loss: 0.10752880573272705
step: 190, loss: 0.1835559606552124
step: 200, loss: 0.05910252407193184
step: 210, loss: 0.04477165266871452
step: 220, loss: 0.12664160132408142
step: 230, loss: 0.1678788661956787
step: 240, loss: 0.05536573380231857
step: 250, loss: 0.09185001999139786
step: 260, loss: 0.24881382286548615
step: 270, loss: 0.34064459800720215
step: 280, loss: 0.08800925314426422
step: 290, loss: 0.08450843393802643
step: 300, loss: 0.04940260574221611
step: 310, loss: 0.17324867844581604
step: 320, loss: 0.08370253443717957
step: 330, loss: 0.07786433398723602
step: 340, loss: 0.24042922258377075
step: 350, loss: 0.15330733358860016
step: 360, loss: 0.05969567224383354
step: 370, loss: 0.0912322923541069
step: 380, loss: 0.06762414425611496
step: 390, loss: 0.04544783756136894
step: 400, loss: 0.09861163794994354
step: 410, loss: 0.28021499514579773
step: 420, loss: 0.21904684603214264
step: 430, loss: 0.048811592161655426
step: 440, loss: 0.0693933293223381
step: 450, loss: 0.0298578180372715
step: 460, loss: 0.02499673143029213
step: 470, loss: 0.08679097890853882
step: 480, loss: 0.11708609014749527
step: 490, loss: 0.05656764656305313
step: 500, loss: 0.014899922534823418
```

```
step: 510, loss: 0.0254064928740263
step: 520, loss: 0.09543817490339279
step: 530, loss: 0.006796353030949831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9326521133302369, f1=0.9295904279797514,
best_f1=0.9295904279797514
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040746476501226425
step: 10, loss: 0.07930020987987518
step: 20, loss: 0.041184552013874054
step: 30, loss: 0.021919498220086098
step: 40, loss: 0.04130446910858154
step: 50, loss: 0.05623244866728783
step: 60, loss: 0.1109478771686554
step: 70, loss: 0.028769657015800476
step: 80, loss: 0.022083386778831482
step: 90, loss: 0.030410176143050194
step: 100, loss: 0.14623090624809265
step: 110, loss: 0.028333930298686028
step: 120, loss: 0.13932687044143677
step: 130, loss: 0.02898227795958519
step: 140, loss: 0.0173685010522604
step: 150, loss: 0.033325839787721634
step: 160, loss: 0.017017504200339317
step: 170, loss: 0.042912695556879044
step: 180, loss: 0.06673379242420197
step: 190, loss: 0.004840028006583452
step: 200, loss: 0.057197000831365585
step: 210, loss: 0.1019350215792656
step: 220, loss: 0.013432099483907223
step: 230, loss: 0.09519903361797333
step: 240, loss: 0.09481914341449738
step: 250, loss: 0.030239000916481018
step: 260, loss: 0.13193847239017487
step: 270, loss: 0.09441884607076645
step: 280, loss: 0.14011994004249573
step: 290, loss: 0.02067827247083187
step: 300, loss: 0.09926552325487137
step: 310, loss: 0.07002530992031097
step: 320, loss: 0.06097737327218056
```

```
step: 330, loss: 0.06777318567037582
step: 340, loss: 0.025087909772992134
step: 350, loss: 0.0031235816422849894
step: 360, loss: 0.06214184686541557
step: 370, loss: 0.040960729122161865
step: 380, loss: 0.23947131633758545
step: 390, loss: 0.01585078239440918
step: 400, loss: 0.082554064691006674
step: 410, loss: 0.017837245017290115
step: 420, loss: 0.1443880945444107
step: 430, loss: 0.12366524338722229
step: 440, loss: 0.006867970805615187
step: 450, loss: 0.012915647588670254
step: 460, loss: 0.07272054255008698
step: 470, loss: 0.05514197424054146
step: 480, loss: 0.008423631079494953
step: 490, loss: 0.12937676906585693
step: 500, loss: 0.02844664826989174
step: 510, loss: 0.021270141005516052
step: 520, loss: 0.23217779397964478
step: 530, loss: 0.04623039811849594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.935933147632312, f1=0.9364055299539169,
best_f1=0.9364055299539169
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11394964903593063
step: 10, loss: 0.04665593057870865
step: 20, loss: 0.015583558939397335
step: 30, loss: 0.10015200078487396
step: 40, loss: 0.044974591583013535
step: 50, loss: 0.028528669849038124
step: 60, loss: 0.03057117573916912
step: 70, loss: 0.01565234363079071
step: 80, loss: 0.14849793910980225
step: 90, loss: 0.0027808041777461767
step: 100, loss: 0.03208616375923157
step: 110, loss: 0.02703363448381424
step: 120, loss: 0.061647795140743256
step: 130, loss: 0.0053372252732515335
step: 140, loss: 0.014097043313086033
```

```
step: 150, loss: 0.012256929650902748
step: 160, loss: 0.04244178161025047
step: 170, loss: 0.02110765315592289
step: 180, loss: 0.0077512082643806934
step: 190, loss: 0.012452003546059132
step: 200, loss: 0.05241933465003967
step: 210, loss: 0.07157687097787857
step: 220, loss: 0.142169788479805
step: 230, loss: 0.024929041042923927
step: 240, loss: 0.020304296165704727
step: 250, loss: 0.1547364592552185
step: 260, loss: 0.14977291226387024
step: 270, loss: 0.021121567115187645
step: 280, loss: 0.03624776378273964
step: 290, loss: 0.059067681431770325
step: 300, loss: 0.14487314224243164
step: 310, loss: 0.1465098261833191
step: 320, loss: 0.021041102707386017
step: 330, loss: 0.011498598381876945
step: 340, loss: 0.012662138789892197
step: 350, loss: 0.09194735437631607
step: 360, loss: 0.03127545490860939
step: 370, loss: 0.07523965835571289
step: 380, loss: 0.043383244425058365
step: 390, loss: 0.015857385471463203
step: 400, loss: 0.05208577588200569
step: 410, loss: 0.0713522806763649
step: 420, loss: 0.017696840688586235
step: 430, loss: 0.03088006190955639
step: 440, loss: 0.2394440770149231
step: 450, loss: 0.030003979802131653
step: 460, loss: 0.06388150900602341
step: 470, loss: 0.0074921417981386185
step: 480, loss: 0.11515750735998154
step: 490, loss: 0.015143980272114277
step: 500, loss: 0.022313840687274933
step: 510, loss: 0.005954783409833908
step: 520, loss: 0.0019257066305726767
step: 530, loss: 0.009686453267931938
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9412313432835822, f1=0.9341983317886932,
best_f1=0.9341983317886932
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024716347455978394
step: 10, loss: 0.002114728558808565
step: 20, loss: 0.11403783410787582
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 30, loss: 0.1081177219748497
step: 40, loss: 0.008218392729759216
step: 50, loss: 0.05374211445450783
step: 60, loss: 0.042963765561580066
step: 70, loss: 0.04858579486608505
step: 80, loss: 0.050894010812044144
step: 90, loss: 0.051157765090465546
step: 100, loss: 0.0034199191723018885
step: 110, loss: 0.12234042584896088
step: 120, loss: 0.006557980086654425
step: 130, loss: 0.05054471269249916
step: 140, loss: 0.03990352526307106
step: 150, loss: 0.004732906818389893
step: 160, loss: 0.01881910301744938
step: 170, loss: 0.041780486702919006
step: 180, loss: 0.06365209817886353
step: 190, loss: 0.021135427057743073
step: 200, loss: 0.014705250971019268
step: 210, loss: 0.0030983020551502705
step: 220, loss: 0.009332027286291122
step: 230, loss: 0.018907373771071434
step: 240, loss: 0.008453295566141605
step: 250, loss: 0.2782549262046814
step: 260, loss: 0.021658392623066902
step: 270, loss: 0.042927686125040054
step: 280, loss: 0.005740853492170572
step: 290, loss: 0.03444804251194
step: 300, loss: 0.007959880866110325
step: 310, loss: 0.006197311915457249
step: 320, loss: 0.11835998296737671
step: 330, loss: 0.02980971895158291
step: 340, loss: 0.020633328706026077
step: 350, loss: 0.15323255956172943
step: 360, loss: 0.09094763547182083
step: 370, loss: 0.00444406596943761
step: 380, loss: 0.008966595865786076
step: 390, loss: 0.0008960357517935336
step: 400, loss: 0.06225397065281868
step: 410, loss: 0.0017729524988681078
step: 420, loss: 0.009174762293696404
step: 430, loss: 0.02441290393471718
```

```
step: 440, loss: 0.002551909303292632
step: 450, loss: 0.05012207105755806
step: 460, loss: 0.011296035721898079
step: 470, loss: 0.003267237450927496
step: 480, loss: 0.017557639628648758
step: 490, loss: 0.04458710551261902
step: 500, loss: 0.05865759402513504
step: 510, loss: 0.08806581795215607
step: 520, loss: 0.027162525802850723
step: 530, loss: 0.16798853874206543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9465364946536494, f1=0.9410125406409661,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014687292277812958
step: 10, loss: 0.063494972884655
step: 20, loss: 0.0031616310589015484
step: 30, loss: 0.00615209573879838
step: 40, loss: 0.0003970696998294443
step: 50, loss: 0.06365698575973511
step: 60, loss: 0.01321143563836813
step: 70, loss: 0.0015704717952758074
step: 80, loss: 0.028307266533374786
step: 90, loss: 0.06075816974043846
step: 100, loss: 0.03154582530260086
step: 110, loss: 0.011924981139600277
step: 120, loss: 0.16917313635349274
step: 130, loss: 0.037356339395046234
step: 140, loss: 0.04048570245504379
step: 150, loss: 0.038906458765268326
step: 160, loss: 0.0033196499571204185
step: 170, loss: 0.1721428781747818
step: 180, loss: 0.00656689191237092
step: 190, loss: 0.0029478988144546747
step: 200, loss: 0.008543316274881363
step: 210, loss: 0.004276036284863949
step: 220, loss: 0.026879042387008667
step: 230, loss: 0.008330823853611946
step: 240, loss: 0.00866883248090744
step: 250, loss: 0.13754773139953613
```

```
step: 260, loss: 0.0006457996787503362
step: 270, loss: 0.001956027001142502
step: 280, loss: 0.007079100701957941
step: 290, loss: 0.0031363849993795156
step: 300, loss: 0.03775187209248543
step: 310, loss: 0.15431049466133118
step: 320, loss: 0.0035404847003519535
step: 330, loss: 0.0010775828268378973
step: 340, loss: 0.0016384883783757687
step: 350, loss: 0.0018092539394274354
step: 360, loss: 0.002707820851355791
step: 370, loss: 0.0014693915145471692
step: 380, loss: 0.000646018423140049
step: 390, loss: 0.000847976712975651
step: 400, loss: 0.002320748521015048
step: 410, loss: 0.06546645611524582
step: 420, loss: 0.17251920700073242
step: 430, loss: 0.02793707698583603
step: 440, loss: 0.0025080896448343992
step: 450, loss: 0.03933923318982124
step: 460, loss: 0.033411331474781036
step: 470, loss: 0.020664727315306664
step: 480, loss: 0.003211691975593567
step: 490, loss: 0.019434867426753044
step: 500, loss: 0.020011812448501587
step: 510, loss: 0.0034404988400638103
step: 520, loss: 0.06614264100790024
step: 530, loss: 0.2079353928565979
/usr/local/lib/python3.7/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=0.9445221445221446, f1=0.9400369003690037,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.046001631766557693
step: 10, loss: 0.0019667113665491343
step: 20, loss: 0.0027580205351114273
step: 30, loss: 0.0017103157006204128
step: 40, loss: 0.09852167218923569
step: 50, loss: 0.0009669801802374423
step: 60, loss: 0.02994876354932785
step: 70, loss: 0.0013609922025352716
step: 80, loss: 0.00038257488631643355
step: 90, loss: 0.016932325437664986
step: 100, loss: 0.017211133614182472
step: 110, loss: 0.0028097983449697495
```

```
step: 120, loss: 0.07682499289512634
step: 130, loss: 0.0021933196112513542
step: 140, loss: 0.0007608943269588053
step: 150, loss: 0.002335913712158799
step: 160, loss: 0.013331961818039417
step: 170, loss: 0.006510373204946518
step: 180, loss: 0.0003424371243454516
step: 190, loss: 0.204850971698761
step: 200, loss: 0.03975479304790497
step: 210, loss: 0.002639623824506998
step: 220, loss: 0.0055303568951785564
step: 230, loss: 0.004119417164474726
step: 240, loss: 0.0007966879638843238
step: 250, loss: 0.025697000324726105
step: 260, loss: 0.061692457646131516
step: 270, loss: 0.0007736345869489014
step: 280, loss: 0.022189009934663773
step: 290, loss: 0.0020685107447206974
step: 300, loss: 0.0010664963629096746
step: 310, loss: 0.17679330706596375
step: 320, loss: 0.00074596336344257
step: 330, loss: 0.0024365244898945093
step: 340, loss: 0.004276445601135492
step: 350, loss: 0.021346498280763626
step: 360, loss: 0.02671041712164879
step: 370, loss: 0.0039004553109407425
step: 380, loss: 0.001398703665472567
step: 390, loss: 0.0007901432691141963
step: 400, loss: 0.0008685392094776034
step: 410, loss: 0.004039010964334011
step: 420, loss: 0.00835799053311348
step: 430, loss: 0.0004476550093386322
step: 440, loss: 0.036592304706573486
step: 450, loss: 0.15391948819160461
step: 460, loss: 0.006266277749091387
step: 470, loss: 0.004649884067475796
step: 480, loss: 0.005309200379997492
step: 490, loss: 0.11735694855451584
step: 500, loss: 0.0016618912341073155
step: 510, loss: 0.15895451605319977
step: 520, loss: 0.0011683361371979117
step: 530, loss: 0.002325385343283415
/usr/local/lib/python3.7/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=0.942008486562942, f1=0.9406819243344232,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032975174486637115
step: 10, loss: 0.01138349063694477
step: 20, loss: 0.0031910748220980167
step: 30, loss: 0.005582951009273529
step: 40, loss: 0.0010575171327218413
step: 50, loss: 0.02313649095594883
step: 60, loss: 0.006805861834436655
step: 70, loss: 0.0007051190477795899
step: 80, loss: 0.009133879095315933
step: 90, loss: 0.001557728392072022
step: 100, loss: 0.019185127690434456
step: 110, loss: 7.811591785866767e-05
step: 120, loss: 0.003030474530532956
step: 130, loss: 6.940739694982767e-05
step: 140, loss: 6.798998947488144e-05
step: 150, loss: 0.0020681428723037243
step: 160, loss: 0.0003391109057702124
step: 170, loss: 0.006118484772741795
step: 180, loss: 0.09472525119781494
step: 190, loss: 0.0008933254284784198
step: 200, loss: 7.427971286233515e-05
step: 210, loss: 0.00010973414464388043
step: 220, loss: 0.0010168836452066898
step: 230, loss: 0.0002748446713667363
step: 240, loss: 0.0029777847230434418
step: 250, loss: 0.0043328325264155865
step: 260, loss: 0.002777281915768981
step: 270, loss: 0.0010150490561500192
step: 280, loss: 0.0026030505541712046
step: 290, loss: 0.010533930733799934
step: 300, loss: 0.00028107574325986207
step: 310, loss: 0.0009988019010052085
step: 320, loss: 0.008554598316550255
step: 330, loss: 0.012142864055931568
step: 340, loss: 0.012992204166948795
step: 350, loss: 0.004198538139462471
step: 360, loss: 0.021540937945246696
step: 370, loss: 0.004868345335125923
step: 380, loss: 0.005452133249491453
step: 390, loss: 0.004257442429661751
step: 400, loss: 0.005437306594103575
step: 410, loss: 0.002577094826847315
step: 420, loss: 0.005304151214659214
step: 430, loss: 0.000780255941208452
step: 440, loss: 0.002403735416010022
step: 450, loss: 0.0011024996638298035
```

```
step: 460, loss: 0.0008739629411138594
step: 470, loss: 0.09996365755796432
step: 480, loss: 0.015473808161914349
step: 490, loss: 0.027942515909671783
step: 500, loss: 0.0013363909674808383
step: 510, loss: 0.00016146311827469617
step: 520, loss: 0.006769634783267975
step: 530, loss: 0.025761298835277557
/usr/local/lib/python3.7/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=0.9443413729128015, f1=0.9404706968158745,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024103871546685696
step: 10, loss: 0.03849535807967186
step: 20, loss: 0.006100665777921677
step: 30, loss: 0.011503932066261768
step: 40, loss: 8.492293272865936e-05
step: 50, loss: 0.004810789134353399
step: 60, loss: 0.0038244009483605623
step: 70, loss: 0.018468840047717094
step: 80, loss: 0.03511032462120056
step: 90, loss: 0.001762346480973065
step: 100, loss: 0.04315508157014847
step: 110, loss: 0.0009582046768628061
step: 120, loss: 0.001416314160451293
step: 130, loss: 0.0001769311638781801
step: 140, loss: 6.697934441035613e-05
step: 150, loss: 0.0005179784493520856
step: 160, loss: 0.0001327722129644826
step: 170, loss: 0.093948595225811
step: 180, loss: 0.0006485946360044181
step: 190, loss: 0.03317508101463318
step: 200, loss: 0.002208047779276967
step: 210, loss: 0.06988739967346191
step: 220, loss: 0.00081581500012604892
step: 230, loss: 0.1250859946012497
step: 240, loss: 0.011914853006601334
step: 250, loss: 0.0010293200612068176
step: 260, loss: 4.6993936848593876e-05
step: 270, loss: 0.10410647094249725
step: 280, loss: 8.39313433971256e-05
step: 290, loss: 0.01225876621901989
step: 300, loss: 0.00016632118786219507
step: 310, loss: 0.005942799616605043
```

```
step: 320, loss: 0.0001541441101543352
step: 330, loss: 0.000769974896684289
step: 340, loss: 0.00826876237988472
step: 350, loss: 0.004147735424339771
step: 360, loss: 0.0181212667375803
step: 370, loss: 0.12901072204113007
step: 380, loss: 0.00013247864262666553
step: 390, loss: 0.03540768846869469
step: 400, loss: 0.0011801145737990737
step: 410, loss: 0.014359957538545132
step: 420, loss: 0.0004852734273299575
step: 430, loss: 0.0003175260208081454
step: 440, loss: 0.03925265744328499
step: 450, loss: 0.00016156751371454448
step: 460, loss: 0.0024891577195376158
step: 470, loss: 0.025097720324993134
step: 480, loss: 0.007695029955357313
step: 490, loss: 0.0007775851408950984
step: 500, loss: 0.002086963038891554
step: 510, loss: 0.0018582235788926482
step: 520, loss: 0.0028901153709739447
step: 530, loss: 0.0010281400755047798
/usr/local/lib/python3.7/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=0.9454209065679925, f1=0.9462068965517241,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007986869313754141
step: 10, loss: 0.009526240639388561
step: 20, loss: 0.00042426559957675636
step: 30, loss: 0.07038363814353943
step: 40, loss: 0.00037149793934077024
step: 50, loss: 0.0006401751888915896
step: 60, loss: 0.0020269523374736311
step: 70, loss: 0.0018265905091539025
step: 80, loss: 0.0013442824129015207
step: 90, loss: 0.032223690301179886
step: 100, loss: 0.0003887044149450958
step: 110, loss: 0.007254039868712425
step: 120, loss: 0.0007171356701292098
step: 130, loss: 0.0001382144691888243
step: 140, loss: 0.00010778483556350693
step: 150, loss: 0.025941750034689903
step: 160, loss: 0.003458930877968669
step: 170, loss: 0.0007641335832886398
```

```
step: 180, loss: 0.00013440771726891398
step: 190, loss: 9.142055205302313e-05
step: 200, loss: 8.677988080307841e-05
step: 210, loss: 0.0033377332147210836
step: 220, loss: 0.000313225609716028
step: 230, loss: 0.0009758905507624149
step: 240, loss: 0.002376803196966648
step: 250, loss: 0.001196251600049436
step: 260, loss: 0.00036878083483316004
step: 270, loss: 0.000841239292640239
step: 280, loss: 0.002405353356152773
step: 290, loss: 0.00044918773346580565
step: 300, loss: 0.02896077372133732
step: 310, loss: 0.04119708761572838
step: 320, loss: 0.037138767540454865
step: 330, loss: 0.00018552615074440837
step: 340, loss: 0.012519050389528275
step: 350, loss: 0.12799939513206482
step: 360, loss: 0.007596786133944988
step: 370, loss: 0.005038033705204725
step: 380, loss: 0.011356065049767494
step: 390, loss: 0.00033351643099449575
step: 400, loss: 0.014245381578803062
step: 410, loss: 0.007099521346390247
step: 420, loss: 0.00037896784488111734
step: 430, loss: 0.016936592757701874
step: 440, loss: 4.354861084721051e-05
step: 450, loss: 0.00020727816445287317
step: 460, loss: 8.294639701489359e-05
step: 470, loss: 0.00022154567705001682
step: 480, loss: 0.016293713822960854
step: 490, loss: 0.00167557829990983
step: 500, loss: 5.265326035441831e-05
step: 510, loss: 0.00025884227943606675
step: 520, loss: 0.00030557953868992627
step: 530, loss: 0.04369933903217316
/usr/local/lib/python3.7/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=0.9435559736594544, f1=0.9443665264142123,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017421870143152773
step: 10, loss: 0.00030060444260016084
step: 20, loss: 0.0116575937718153
step: 30, loss: 0.0009094635606743395
```

```
step: 40, loss: 8.752760186325759e-05
step: 50, loss: 0.0050322446040809155
step: 60, loss: 0.0023162902798503637
step: 70, loss: 7.742394518572837e-05
step: 80, loss: 7.083051605150104e-05
step: 90, loss: 0.00033174495911225677
step: 100, loss: 7.481210195692256e-05
step: 110, loss: 0.013241109438240528
step: 120, loss: 4.185739089734852e-05
step: 130, loss: 0.00012617031461559236
step: 140, loss: 0.000279762054560706
step: 150, loss: 8.921947301132604e-05
step: 160, loss: 0.0603373721241951
step: 170, loss: 0.02212977036833763
step: 180, loss: 0.024391544982790947
step: 190, loss: 0.0001011011190712452
step: 200, loss: 0.00017506058793514967
step: 210, loss: 0.010403911583125591
step: 220, loss: 0.0007629717001691461
step: 230, loss: 0.0013323896564543247
step: 240, loss: 0.05178608000278473
step: 250, loss: 0.00015502434689551592
step: 260, loss: 0.000264617643551901
step: 270, loss: 0.0032069033477455378
step: 280, loss: 0.04715064540505409
step: 290, loss: 6.743278936482966e-05
step: 300, loss: 0.004140723496675491
step: 310, loss: 0.032793302088975906
step: 320, loss: 0.0020260068122297525
step: 330, loss: 8.794397581368685e-05
step: 340, loss: 3.633850792539306e-05
step: 350, loss: 0.0023486220743507147
step: 360, loss: 7.090158760547638e-05
step: 370, loss: 0.003066703211516142
step: 380, loss: 0.0005142097943462431
step: 390, loss: 8.20190689410083e-05
step: 400, loss: 0.0010659538675099611
step: 410, loss: 0.00041596413939259946
step: 420, loss: 6.707227294100448e-05
step: 430, loss: 5.999591667205095e-05
step: 440, loss: 0.0001664972078287974
step: 450, loss: 8.200949378078803e-05
step: 460, loss: 7.05350685166195e-05
step: 470, loss: 0.00014323422510642558
step: 480, loss: 0.0021771294996142387
step: 490, loss: 0.014534315094351768
step: 500, loss: 0.014138770289719105
step: 510, loss: 0.0814441484212875
```

```
step: 520, loss: 0.014706892892718315
step: 530, loss: 0.019472217187285423
/usr/local/lib/python3.7/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=0.9434137291280148, f1=0.9424758398527383,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014280207687988877
step: 10, loss: 0.054928749799728394
step: 20, loss: 0.0036354502663001632
step: 30, loss: 0.00021496917179320008
step: 40, loss: 0.00011865428677992895
step: 50, loss: 9.790756303118542e-05
step: 60, loss: 0.00025653516058810055
step: 70, loss: 0.0028725366573780775
step: 80, loss: 0.03627188876271248
step: 90, loss: 0.003485277993604541
step: 100, loss: 0.0055272141471505165
step: 110, loss: 0.00010613568883854896
step: 120, loss: 0.0009102553594857454
step: 130, loss: 0.00016344775212928653
step: 140, loss: 0.00014640232257079333
step: 150, loss: 0.0003646143595688045
step: 160, loss: 0.0062974924221634865
step: 170, loss: 0.002637010533362627
step: 180, loss: 0.0012219416676089168
step: 190, loss: 0.002067738911136985
step: 200, loss: 0.00014443544205278158
step: 210, loss: 0.0006175670423544943
step: 220, loss: 0.03915534168481827
step: 230, loss: 0.00013024611689615995
step: 240, loss: 0.005880070850253105
step: 250, loss: 0.0038486295379698277
step: 260, loss: 0.00047629053005948663
step: 270, loss: 0.00019306770991533995
step: 280, loss: 0.00026336737209931016
step: 290, loss: 0.0003046024357900023
step: 300, loss: 0.00021649443078786135
step: 310, loss: 0.2258247584104538
step: 320, loss: 0.0032443893142044544
step: 330, loss: 0.00018231300055049360
step: 340, loss: 0.03249190375208855
step: 350, loss: 0.0001842075289459899
step: 360, loss: 0.00041369657265022397
step: 370, loss: 0.0010958024067804217
```

```
step: 380, loss: 0.0003853289526887238
step: 390, loss: 0.004779044538736343
step: 400, loss: 0.000923917512409389
step: 410, loss: 6.616055907215923e-05
step: 420, loss: 0.19661830365657806
step: 430, loss: 0.007328664883971214
step: 440, loss: 0.0012116103898733854
step: 450, loss: 0.00013884727377444506
step: 460, loss: 0.0006854557432234287
step: 470, loss: 0.0004468016268219799
step: 480, loss: 0.024553198367357254
step: 490, loss: 0.008431723341345787
step: 500, loss: 0.0022819912992417812
step: 510, loss: 0.0020354490261524916
step: 520, loss: 0.0012864762684330344
step: 530, loss: 0.03363858908414841
/usr/local/lib/python3.7/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=0.945287356321839, f1=0.9452054794520548,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012800995260477066
step: 10, loss: 0.007397601380944252
step: 20, loss: 0.011334016919136047
step: 30, loss: 0.0008189913351088762
step: 40, loss: 0.0006309784366749227
step: 50, loss: 8.77908751135692e-05
step: 60, loss: 9.946687350748107e-05
step: 70, loss: 0.00044517763308249414
step: 80, loss: 0.00014696104335598648
step: 90, loss: 7.892821304267272e-05
step: 100, loss: 0.25339099764823914
step: 110, loss: 0.000264665053691715
step: 120, loss: 0.00021672299772203533
step: 130, loss: 0.0002779737988021225
step: 140, loss: 0.0001931258011609316
step: 150, loss: 0.0001405411312589422
step: 160, loss: 0.23229104280471802
step: 170, loss: 0.007582851219922304
step: 180, loss: 0.00017243098409380764
step: 190, loss: 0.006866671144962311
step: 200, loss: 0.009325525723397732
step: 210, loss: 0.0005115412641316652
step: 220, loss: 0.00014909128367435187
step: 230, loss: 0.0004845480143558234
```

```
step: 240, loss: 0.0005438024527393281
step: 250, loss: 9.228655835613608e-05
step: 260, loss: 0.001828873180784285
step: 270, loss: 0.038475509732961655
step: 280, loss: 0.000312718068016693
step: 290, loss: 0.0015613174764439464
step: 300, loss: 0.00015098198491614312
step: 310, loss: 0.02290685474872589
step: 320, loss: 0.0007166791474446654
step: 330, loss: 0.0013518014457076788
step: 340, loss: 0.004671149887144566
step: 350, loss: 6.198760820552707e-05
step: 360, loss: 8.889666059985757e-05
step: 370, loss: 0.0003195870667695999
step: 380, loss: 0.0004328649374656379
step: 390, loss: 0.0007689158082939684
step: 400, loss: 8.61437147250399e-05
step: 410, loss: 0.001363467308692634
step: 420, loss: 6.110304821049795e-05
step: 430, loss: 0.00030681968200951815
step: 440, loss: 0.04392363503575325
step: 450, loss: 0.0011519375257194042
step: 460, loss: 8.322561188833788e-05
step: 470, loss: 0.00029840407660230994
step: 480, loss: 0.000653691531624645
step: 490, loss: 8.139598503476009e-05
step: 500, loss: 7.32009721104987e-05
step: 510, loss: 0.0010024715447798371
step: 520, loss: 0.0042485641315579414
step: 530, loss: 0.00018363777780905366
/usr/local/lib/python3.7/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=0.9422263973696571, f1=0.9431658055425082,
best_f1=0.9410125406409661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.960732182255015e-05
step: 10, loss: 0.00042999707511626184
step: 20, loss: 0.00014793244190514088
step: 30, loss: 6.25775137450546e-05
step: 40, loss: 0.0012424177257344127
step: 50, loss: 0.0012465803883969784
step: 60, loss: 7.61683113523759e-05
step: 70, loss: 0.00018565767095424235
step: 80, loss: 8.942541171563789e-05
step: 90, loss: 8.69371069711633e-05
```

```
step: 100, loss: 0.0005056869122199714
step: 110, loss: 4.098471617908217e-05
step: 120, loss: 8.696525765117258e-05
step: 130, loss: 4.2085466702701524e-05
step: 140, loss: 0.00018276686023455113
step: 150, loss: 0.00013504074013326317
step: 160, loss: 0.00021888813353143632
step: 170, loss: 6.049140574759804e-05
step: 180, loss: 0.00010992607712978497
step: 190, loss: 0.00022860814351588488
step: 200, loss: 0.00017557202954776585
step: 210, loss: 0.0003147621173411608
step: 220, loss: 6.298694643191993e-05
step: 230, loss: 0.0015872414223849773
step: 240, loss: 8.17413383629173e-05
step: 250, loss: 0.00020895901252515614
step: 260, loss: 7.892778376117349e-05
step: 270, loss: 0.00010269500489812344
step: 280, loss: 6.393322837539017e-05
step: 290, loss: 7.245936285471544e-05
step: 300, loss: 0.00011562016152311116
step: 310, loss: 8.498979877913371e-05
step: 320, loss: 7.733840902801603e-05
step: 330, loss: 0.00010245312296319753
step: 340, loss: 0.00013509226846508682
step: 350, loss: 5.028725718148053e-05
step: 360, loss: 0.00945375021547079
step: 370, loss: 0.00018271345470566303
step: 380, loss: 0.00012359167158138007
step: 390, loss: 9.928361396305263e-05
step: 400, loss: 0.00012910460645798594
step: 410, loss: 0.00016628281446173787
step: 420, loss: 0.00016343765310011804
step: 430, loss: 0.0025610842276364565
step: 440, loss: 7.519911014242098e-05
step: 450, loss: 0.00012422654253896326
step: 460, loss: 0.00020301582117099315
step: 470, loss: 4.251319478498772e-05
step: 480, loss: 3.5872777516487986e-05
step: 490, loss: 0.00010354393452871591
step: 500, loss: 0.0003578302275855094
step: 510, loss: 0.00010154621850233525
step: 520, loss: 6.344253779388964e-05
step: 530, loss: 7.708761404501274e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9482517482517483, f1=0.9439555349698935,
best_f1=0.9439555349698935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001554416521685198
step: 10, loss: 0.00028619752265512943
step: 20, loss: 6.741427932865918e-05
step: 30, loss: 8.692896517459303e-05
step: 40, loss: 7.110050501069054e-05
step: 50, loss: 0.0001951863814610988
step: 60, loss: 8.810640429146588e-05
step: 70, loss: 0.00013508237316273153
step: 80, loss: 0.0014446849236264825
step: 90, loss: 5.3844447393203154e-05
step: 100, loss: 6.712056347168982e-05
step: 110, loss: 0.00010322307207388803
step: 120, loss: 7.969005673658103e-05
step: 130, loss: 9.796745143830776e-05
step: 140, loss: 8.754291775403544e-05
step: 150, loss: 0.00012458587298169732
step: 160, loss: 9.012393275042996e-05
step: 170, loss: 0.0001316552807111293
step: 180, loss: 0.000153371220221743
step: 190, loss: 0.00027530782972462475
step: 200, loss: 0.00011644933692878112
step: 210, loss: 0.0002049535105470568
step: 220, loss: 4.835866275243461e-05
step: 230, loss: 9.087628859560937e-05
step: 240, loss: 5.212984251556918e-05
step: 250, loss: 0.00011081374395871535
step: 260, loss: 7.476411701645702e-05
step: 270, loss: 0.00014739613106939942
step: 280, loss: 0.00013805096386931837
step: 290, loss: 5.02733928442467e-05
step: 300, loss: 0.00011784651724155992
step: 310, loss: 6.232441228348762e-05
step: 320, loss: 9.295075869886205e-05
step: 330, loss: 9.018988203024492e-05
step: 340, loss: 7.638589158887044e-05
step: 350, loss: 6.574390863534063e-05
step: 360, loss: 0.00010585584095679224
step: 370, loss: 0.0003997193416580558
step: 380, loss: 0.00016172138566616923
step: 390, loss: 0.015425801277160645

```
step: 400, loss: 0.0024498451966792345
step: 410, loss: 3.320696123410016e-05
step: 420, loss: 0.0001716708211461082
step: 430, loss: 0.0002121241413988173
step: 440, loss: 7.957580965012312e-05
step: 450, loss: 0.018593259155750275
step: 460, loss: 0.1174052432179451
step: 470, loss: 0.00010679865954443812
step: 480, loss: 0.0004513929598033428
step: 490, loss: 0.0006056393031030893
step: 500, loss: 0.00025377864949405193
step: 510, loss: 0.013071910478174686
step: 520, loss: 3.756830483325757e-05
step: 530, loss: 8.190731750801206e-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=0.9492314857941314, f1=0.9435185185185185,
best_f1=0.9435185185185185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.56557604088448e-05
step: 10, loss: 0.00011170264042448252
step: 20, loss: 9.702180250314996e-05
step: 30, loss: 0.01995258964598179
step: 40, loss: 8.083878492470831e-05
step: 50, loss: 7.671866478631273e-05
step: 60, loss: 0.0001820346078602597
step: 70, loss: 5.541603240999393e-05
step: 80, loss: 0.0001151569013018161
step: 90, loss: 4.9097128794528544e-05
step: 100, loss: 3.393782753846608e-05
step: 110, loss: 0.0003974122228100896
step: 120, loss: 4.4020547647960484e-05
step: 130, loss: 4.869049007538706e-05
step: 140, loss: 6.517746078316122e-05
step: 150, loss: 0.00010446574742672965
step: 160, loss: 0.00010828048107214272
step: 170, loss: 3.866981569444761e-05
step: 180, loss: 0.00011905882274731994
step: 190, loss: 0.0002271891717100516
step: 200, loss: 7.312930392799899e-05
step: 210, loss: 3.2167547033168375e-05
```

```
step: 220, loss: 7.960476796142757e-05
step: 230, loss: 9.635018068365753e-05
step: 240, loss: 6.605796806979924e-05
step: 250, loss: 4.0734117646934465e-05
step: 260, loss: 4.1927163692889735e-05
step: 270, loss: 7.47441517887637e-05
step: 280, loss: 4.493645246839151e-05
step: 290, loss: 9.504274203209206e-05
step: 300, loss: 7.676041423110291e-05
step: 310, loss: 0.0003904491022694856
step: 320, loss: 0.00010162662510992959
step: 330, loss: 0.00011216444545425475
step: 340, loss: 0.00010194830247201025
step: 350, loss: 0.002842389978468418
step: 360, loss: 0.00020386552205309272
step: 370, loss: 0.00020432549354154617
step: 380, loss: 9.524673805572093e-05
step: 390, loss: 9.552423580316827e-05
step: 400, loss: 0.02069278247654438
step: 410, loss: 7.593898044433445e-05
step: 420, loss: 5.1202707254560664e-05
step: 430, loss: 3.1020470487419516e-05
step: 440, loss: 5.408784272731282e-05
step: 450, loss: 0.00012029504432575777
step: 460, loss: 7.539407670265064e-05
step: 470, loss: 7.190405449364334e-05
step: 480, loss: 8.961065032053739e-05
step: 490, loss: 0.0020498994272202253
step: 500, loss: 7.26745420251973e-05
step: 510, loss: 6.828025652794167e-05
step: 520, loss: 8.354450255865231e-05
step: 530, loss: 5.331663851393387e-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 15: dev_f1=0.9502093997208004, f1=0.9429763560500695,
best_f1=0.9429763560500695
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 [00:32, 174.82it/s]
load_f1 = 0.9481000926784059
real_f1 = 0.9487418452935694
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:29, 149.67it/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 488 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
```

```
Requirement already satisfied: 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: 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: 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: 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: 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: 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: 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: 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: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
```

```
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: 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: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) … done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=d7cccdcbc5374ca024b15ef6adfce39a4d5467b4108aff8c0a76bf32259038d6
  Stored in directory: /tmp/pip-ephem-wheel-cache-5xylsv77/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

```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 roberta \
  --fp16 \
  --save_model \
  --dk general  \
  --dk product \
  --summarize
```

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

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

```
Selected optimization level 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: 0.4516790807247162
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.4210526315789474, f1=0.37681159420289856,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4609569311141968
/usr/local/lib/python3.7/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=0.3880597014925373, f1=0.3466666666666666,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4586017429828644
/usr/local/lib/python3.7/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=0.37333333333333335, f1=0.3170731707317073,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30387258529663086
/usr/local/lib/python3.7/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=0.2857142857142857, f1=0.2745098039215686,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36570069193840027
/usr/local/lib/python3.7/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=0.28, f1=0.2692307692307693, best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28344348073005676
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.34146341463414637, f1=0.3023255813953489,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6264019012451172
/usr/local/lib/python3.7/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=0.4057971014492754, f1=0.3561643835616438,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5302332639694214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5, f1=0.36065573770491804, best_f1=0.36065573770491804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33577069640159607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5333333333333333, f1=0.4444444444444444,
best_f1=0.4444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.47006282210350037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6428571428571429, f1=0.4864864864864865,
best_f1=0.4864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34968385100364685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.689655172413793, f1=0.5161290322580646,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20741282403469086
/usr/local/lib/python3.7/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=0.689655172413793, f1=0.5517241379310344,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24317359924316406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7407407407407408, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.09043914079666138
/usr/local/lib/python3.7/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=0.7407407407407408, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21179020404815674
/usr/local/lib/python3.7/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=0.7407407407407408, f1=0.5925925925925927,
best_f1=0.5925925925925927
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, 106406.93it/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 = 0.689655172413793
real_f1 = 0.689655172413793
```

```
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.82it/s]
```

3.1.2 DBLP-ACM - Running the matcher

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


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

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

```
Downloading: 100% 481/481 [00:00<00:00, 500kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 34.3MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 30.2MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 68.4MB/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.")
step: 0, loss: 0.6397694945335388
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5084032416343689
step: 20, loss: 0.4772605895996094
step: 30, loss: 0.2827351689338684
step: 40, loss: 0.35578182339668274
step: 50, loss: 0.6614904999732971
step: 60, loss: 0.4788443148136139
step: 70, loss: 0.40732985734939575
step: 80, loss: 0.5828161835670471
step: 90, loss: 0.4332062005996704
step: 100, loss: 0.45420804619789124
step: 110, loss: 0.6317482590675354
step: 120, loss: 0.47220805287361145
step: 130, loss: 0.33304086327552795
step: 140, loss: 0.39843225479125977
step: 150, loss: 0.5953852534294128
step: 160, loss: 0.5303943157196045
step: 170, loss: 0.47353503108024597
step: 180, loss: 0.40128093957901
step: 190, loss: 0.2273622304201126
step: 200, loss: 0.13679052889347076
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
```

```
step: 210, loss: 0.5517314672470093
step: 220, loss: 0.24789053201675415
step: 230, loss: 0.1862126886844635
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.886021505376344, f1=0.8932461873638344,
best_f1=0.8932461873638344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06438636779785156
step: 10, loss: 0.19339360296726227
step: 20, loss: 0.05729708820581436
step: 30, loss: 0.16120372712612152
step: 40, loss: 0.05063290521502495
step: 50, loss: 0.021420473232865334
step: 60, loss: 0.009079274721443653
step: 70, loss: 0.048561304807662964
step: 80, loss: 0.0033870262559503317
step: 90, loss: 0.007828092202544212
step: 100, loss: 0.0047811768017709255
step: 110, loss: 0.15304724872112274
step: 120, loss: 0.040958844125270844
step: 130, loss: 0.07399177551269531
step: 140, loss: 0.018196213990449905
step: 150, loss: 0.15293078124523163
step: 160, loss: 0.020596317946910858
step: 170, loss: 0.0040773265063762665
step: 180, loss: 0.008599843829870224
step: 190, loss: 0.007781761698424816
step: 200, loss: 0.01101361121982336
step: 210, loss: 0.059571441262960434
step: 220, loss: 0.008632175624370575
step: 230, loss: 0.0014350101118907332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.9751131221719457,
best_f1=0.9751131221719457
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033235255628824234
step: 10, loss: 0.016752339899539948
step: 20, loss: 0.013631905429065228
step: 30, loss: 0.0008857263019308448
step: 40, loss: 0.012284415774047375
step: 50, loss: 0.008028370328247547
step: 60, loss: 0.00942761916667223
step: 70, loss: 0.011119278147816658
step: 80, loss: 0.027560045942664146
step: 90, loss: 0.006444413680583239
step: 100, loss: 0.006185253616422415
step: 110, loss: 0.008997339755296707
step: 120, loss: 0.006318960804492235
step: 130, loss: 0.005917144473642111
step: 140, loss: 0.0011754361912608147
step: 150, loss: 0.06015288457274437
step: 160, loss: 0.015666360035538673
step: 170, loss: 0.010457859374582767
step: 180, loss: 0.021351605653762817
step: 190, loss: 0.04763951152563095
step: 200, loss: 0.013486607931554317
step: 210, loss: 0.0019518263870850205
step: 220, loss: 0.0037291203625500202
step: 230, loss: 0.005441464018076658
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9754464285714286,
best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009919722564518452
step: 10, loss: 0.002000285079702735
step: 20, loss: 0.0025424957275390625
step: 30, loss: 0.0005642788601107895
step: 40, loss: 0.07706528156995773
step: 50, loss: 0.01552465558052063
step: 60, loss: 0.0011030000168830156
step: 70, loss: 0.0023540782276540995
step: 80, loss: 0.0027692054864019156
step: 90, loss: 0.002400730038061738
step: 100, loss: 0.0030883746221661568
step: 110, loss: 0.0017426552949473262
step: 120, loss: 0.05721547454595566
```

```
step: 130, loss: 0.04018848016858101
step: 140, loss: 0.0036523137241601944
step: 150, loss: 0.0007424487848766148
step: 160, loss: 0.06077278032898903
step: 170, loss: 0.0540262870490551
step: 180, loss: 0.16570539772510529
step: 190, loss: 0.00304606044664979
step: 200, loss: 0.00556445587426424
step: 210, loss: 0.08960946649312973
step: 220, loss: 0.0007346186321228743
step: 230, loss: 0.0016074070008471608
/usr/local/lib/python3.7/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=0.9712389380530975, f1=0.9787709497206705,
best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05638273060321808
step: 10, loss: 0.003943208605051041
step: 20, loss: 0.034409694373607635
step: 30, loss: 0.0007063202792778611
step: 40, loss: 0.001999777043238282
step: 50, loss: 0.003639088710770011
step: 60, loss: 0.0017282793996855617
step: 70, loss: 0.008450284600257874
step: 80, loss: 0.04858442023396492
step: 90, loss: 0.0833168625831604
step: 100, loss: 0.0007532487506978214
step: 110, loss: 0.0011638260912150145
step: 120, loss: 0.0015850345371291041
step: 130, loss: 0.0010227750753983855
step: 140, loss: 0.0035785213112831116
step: 150, loss: 0.04849456995725632
step: 160, loss: 0.0015685491962358356
step: 170, loss: 0.004342665430158377
step: 180, loss: 0.031145259737968445
step: 190, loss: 0.011216496117413044
step: 200, loss: 0.002069701673462987
step: 210, loss: 0.0004222589486744255
step: 220, loss: 0.0004476374597288668
step: 230, loss: 0.005684445612132549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037625785917043686
step: 10, loss: 0.003870048327371478
step: 20, loss: 0.0017778188921511173
step: 30, loss: 0.004155496601015329
step: 40, loss: 0.00037646127748303115
step: 50, loss: 0.0005714840954169631
step: 60, loss: 0.0021685101091861725
step: 70, loss: 0.0039964765310287476
step: 80, loss: 0.0031543460208922625
step: 90, loss: 0.0004220866831019521
step: 100, loss: 0.0004723835445474833
step: 110, loss: 0.09808791428804398
step: 120, loss: 0.00022688109311275184
step: 130, loss: 0.002333609852939844
step: 140, loss: 0.0011471961624920368
step: 150, loss: 0.00021268369164317846
step: 160, loss: 0.0011930419132113457
step: 170, loss: 0.0013093315064907074
step: 180, loss: 0.0003634998865891248
step: 190, loss: 0.0007414400461129844
step: 200, loss: 0.01816507987678051
step: 210, loss: 0.005332258529961109
step: 220, loss: 0.0009192571742460132
step: 230, loss: 0.0029992840718477964
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887133182844244, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038247834891080856
step: 10, loss: 0.0011907409643754363
step: 20, loss: 0.0007839194731786847
step: 30, loss: 0.0007631017942912877
step: 40, loss: 0.0017503555864095688
```

```
step: 50, loss: 0.0031751859933137894
step: 60, loss: 0.0024785317946225405
step: 70, loss: 0.0008009775774553418
step: 80, loss: 0.0005505466251634061
step: 90, loss: 0.049262754461912155
step: 100, loss: 0.0007260781130753458
step: 110, loss: 0.0003306999569758773
step: 120, loss: 0.0010147620923817158
step: 130, loss: 0.00428803451359272
step: 140, loss: 0.0007945471443235874
step: 150, loss: 0.007325616665184498
step: 160, loss: 0.000537489599082619
step: 170, loss: 0.0006887627532705665
step: 180, loss: 0.0001936090993694961
step: 190, loss: 0.0004054440651088953
step: 200, loss: 0.0006446211482398212
step: 210, loss: 0.0003932759864255786
step: 220, loss: 0.0006063197506591678
step: 230, loss: 0.0009417465189471841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898762654668166, f1=0.9785310734463276,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016283217119053006
step: 10, loss: 0.004678468685597181
step: 20, loss: 0.002311501419171691
step: 30, loss: 0.001195794902741909
step: 40, loss: 0.001371339662000537
step: 50, loss: 0.005418069660663605
step: 60, loss: 0.0005352755542844534
step: 70, loss: 0.00017719810421112925
step: 80, loss: 0.0009299312368966639
step: 90, loss: 0.0009240771178156137
step: 100, loss: 0.007212469354271889
step: 110, loss: 0.00048610809608362615
step: 120, loss: 0.0005706362426280975
step: 130, loss: 0.0030510893557220697
step: 140, loss: 0.001699085347354412
step: 150, loss: 0.18813754618167877
step: 160, loss: 0.0012530875392258167
```

```
step: 170, loss: 0.027197768911719322
step: 180, loss: 0.002895200625061989
step: 190, loss: 0.0004804682976100594
step: 200, loss: 0.0028567006811499596
step: 210, loss: 0.00472351536154747
step: 220, loss: 0.0021510517690330744
step: 230, loss: 0.0017866905545815825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910112359550561, 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: 0.0005121428403072059
step: 10, loss: 0.00042953132651746273
step: 20, loss: 0.0003430380311328918
step: 30, loss: 0.0004896406317129731
step: 40, loss: 0.00044577609514817595
step: 50, loss: 0.0006898912834003568
step: 60, loss: 0.002086818916723132
step: 70, loss: 0.08164200186729431
step: 80, loss: 0.00023035988851916045
step: 90, loss: 0.046604324132204056
step: 100, loss: 0.0001594762143213302
step: 110, loss: 0.00010415390715934336
step: 120, loss: 0.00028673611814156175
step: 130, loss: 0.0003647439007181674
step: 140, loss: 0.00017784637748263776
step: 150, loss: 0.002383199753239751
step: 160, loss: 0.00024894767557270825
step: 170, loss: 7.011480192886665e-05
step: 180, loss: 0.0001333811233052984
step: 190, loss: 5.219010927248746e-05
step: 200, loss: 0.00015795789659023285
step: 210, loss: 0.005957202985882759
step: 220, loss: 0.0002150252548744902
step: 230, loss: 0.000817024614661932
/usr/local/lib/python3.7/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=0.9830890642615557, f1=0.9785310734463276,
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: 0.00011784752859966683
step: 10, loss: 0.0002899050014093518
step: 20, loss: 0.00022263037681113929
step: 30, loss: 9.750846220413223e-05
step: 40, loss: 0.0007829976384527981
step: 50, loss: 4.168459781794809e-05
step: 60, loss: 0.00029598019318655133
step: 70, loss: 0.0010382015025243163
step: 80, loss: 0.00017769278201740235
step: 90, loss: 0.00016386208881158382
step: 100, loss: 8.594742394052446e-05
step: 110, loss: 0.00020468120055738837
step: 120, loss: 0.00013525254325941205
step: 130, loss: 0.00024111477250698954
step: 140, loss: 0.000391456502256915
step: 150, loss: 0.0004581122484523803
step: 160, loss: 0.00011493014608277008
step: 170, loss: 0.00023963299463503063
step: 180, loss: 0.0003391598293092102
step: 190, loss: 0.00016242408310063183
step: 200, loss: 0.00020829311688430607
step: 210, loss: 0.019248750060796738
step: 220, loss: 0.023169519379734993
step: 230, loss: 0.00019836910360027105
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9841269841269841,
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: 0.00010293570812791586
step: 10, loss: 0.0001578652736498043
step: 20, loss: 0.00024377451336476952
step: 30, loss: 0.00014071079203858972
step: 40, loss: 5.495801451615989e-05
step: 50, loss: 0.00012113696720916778
step: 60, loss: 0.00858637411147356
step: 70, loss: 0.00012823530414607376
step: 80, loss: 0.003100774949416518
step: 90, loss: 0.007127717602998018
step: 100, loss: 0.0014637801796197891
step: 110, loss: 0.14947129786014557
step: 120, loss: 0.0015112125547602773
```

```
step: 130, loss: 0.0003122453053947538
step: 140, loss: 0.029549745842814445
step: 150, loss: 0.0006534033454954624
step: 160, loss: 0.005037821829319
step: 170, loss: 0.0036538001149892807
step: 180, loss: 0.00014258916780818254
step: 190, loss: 9.399164991918951e-05
step: 200, loss: 0.0021617463789880276
step: 210, loss: 0.0007926546386443079
step: 220, loss: 0.0009506108472123742
step: 230, loss: 0.0005367810372263193
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9819413092550789,
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: 0.0009230090654455125
step: 10, loss: 0.000490086677018553
step: 20, loss: 0.004881934262812138
step: 30, loss: 0.04350524768233299
step: 40, loss: 0.0009764126734808087
step: 50, loss: 0.002407512627542019
step: 60, loss: 0.0005322261131368577
step: 70, loss: 0.00045976176625111766
step: 80, loss: 0.0003646301629487425
step: 90, loss: 0.0011044017737731338
step: 100, loss: 0.0003012583183590323
step: 110, loss: 0.0010313987731933594
step: 120, loss: 0.00010816172289196402
step: 130, loss: 0.0001302826130995527
step: 140, loss: 0.00042518184636719525
step: 150, loss: 0.0006162165082059801
step: 160, loss: 0.0006530944956466556
step: 170, loss: 0.000801139569375664
step: 180, loss: 0.0003303303674329072
step: 190, loss: 0.0027439638506621122
step: 200, loss: 0.00012082515604561195
step: 210, loss: 0.000764387019444257
step: 220, loss: 0.01157979667186737
step: 230, loss: 0.0006899043801240623
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9831271091113611,
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: 0.00022096092288848013
step: 10, loss: 0.00013510006829164922
step: 20, loss: 0.0010460158810019493
step: 30, loss: 0.00031751187634654343
step: 40, loss: 0.0016073094448074698
step: 50, loss: 0.0034435035195201635
step: 60, loss: 0.00015798771346453577
step: 70, loss: 6.988779205130413e-05
step: 80, loss: 0.0035867542028427124
step: 90, loss: 0.00012092893302906305
step: 100, loss: 0.0002880466927308589
step: 110, loss: 0.00031176727497950196
step: 120, loss: 0.00021120630844961852
step: 130, loss: 0.00010935833415715024
step: 140, loss: 8.03604707471095e-05
step: 150, loss: 4.1304167098132893e-05
step: 160, loss: 0.0001194788928842172
step: 170, loss: 0.00010162151011172682
step: 180, loss: 0.002662434009835124
step: 190, loss: 0.00012825422163587064
step: 200, loss: 2.0715815480798483e-05
step: 210, loss: 0.00044150411849841475
step: 220, loss: 7.800530875101686e-05
step: 230, loss: 0.00018792458286043257
/usr/local/lib/python3.7/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=0.9863636363636363, f1=0.9794988610478361,
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: 3.358488902449608e-05
step: 10, loss: 4.503105810726993e-05
step: 20, loss: 6.011861842125654e-05
step: 30, loss: 0.00016822910401970148
step: 40, loss: 0.0001160733008873649
step: 50, loss: 4.662191349780187e-05
step: 60, loss: 6.0845475672977045e-05
step: 70, loss: 5.40361215582612395e-05
step: 80, loss: 0.00010444469808135182
step: 90, loss: 0.0004465760721359402
step: 100, loss: 0.00014512201596517116
step: 110, loss: 7.816257857484743e-05
step: 120, loss: 2.1274350729072466e-05
```

```
step: 130, loss: 0.0002563587040640414
step: 140, loss: 5.82879401918035e-05
step: 150, loss: 0.0019540502689778805
step: 160, loss: 0.00020061984832864255
step: 170, loss: 9.477737330598757e-05
step: 180, loss: 3.521699909470044e-05
step: 190, loss: 0.0002543214650359005
step: 200, loss: 0.0003676394990179688
step: 210, loss: 7.770895899739116e-05
step: 220, loss: 0.00017006257257889956
step: 230, loss: 3.203190499334596e-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=0.9887133182844244, f1=0.9830890642615557,
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: 0.0040483297780156136
step: 10, loss: 4.795453423867002e-05
step: 20, loss: 0.0002870052703656256
step: 30, loss: 4.292350058676675e-05
step: 40, loss: 2.548737575125415e-05
step: 50, loss: 4.0114118746714666e-05
step: 60, loss: 0.001409839722327888
step: 70, loss: 0.0073389047756791115
step: 80, loss: 5.113482984597795e-05
step: 90, loss: 5.897847222513519e-05
step: 100, loss: 2.4582275727880187e-05
step: 110, loss: 5.1565304602263495e-05
step: 120, loss: 0.062903150091609955
step: 130, loss: 3.398778062546626e-05
step: 140, loss: 0.0020011630840599537
step: 150, loss: 9.778135427040979e-05
step: 160, loss: 0.04367145523428917
step: 170, loss: 3.828254557447508e-05
step: 180, loss: 6.419179408112541e-05
step: 190, loss: 0.0015568750677630305
step: 200, loss: 7.501921209041029e-05
step: 210, loss: 0.03899999335408211
step: 220, loss: 3.5767654480878264e-05
step: 230, loss: 0.00013752664381172508
/usr/local/lib/python3.7/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=0.9796839729119639,
best_f1=0.9741282339707535
```

```
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, 215.99it/s]
load_f1 = 0.9864864864864865
real_f1 = 0.9820627802690582
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.64it/s]
```

3.1.3 DBLP-GoogleScholar - Running the matcher

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

```
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general  \
  --dk product \
  --summarize


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

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

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

Defaults for this optimization level are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly

```
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6250618100166321
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.47026365995407104
step: 20, loss: 0.36641454696655273
step: 30, loss: 0.4063195586204529
step: 40, loss: 0.39173761010169983
step: 50, loss: 0.672453761100769
step: 60, loss: 0.350340336561203
step: 70, loss: 0.491411030292511
step: 80, loss: 0.26739659905433655
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.18407319486141205
step: 100, loss: 0.384347140789032
step: 110, loss: 0.1467682421207428
step: 120, loss: 0.3680344820022583
step: 130, loss: 0.30079320073127747
step: 140, loss: 0.2035517394542694
step: 150, loss: 0.12020575255155563
step: 160, loss: 0.12821312248706818
step: 170, loss: 0.07061174511909485
step: 180, loss: 0.09533532708883286
step: 190, loss: 0.15929661691188812
step: 200, loss: 0.05828431621193886
step: 210, loss: 0.08984506875276566
step: 220, loss: 0.10575441271066666
step: 230, loss: 0.16241301596164703
step: 240, loss: 0.05382458493113518
step: 250, loss: 0.16650018095970154
step: 260, loss: 0.2623142600059509
step: 270, loss: 0.29540663957595825
step: 280, loss: 0.14070989191532135
step: 290, loss: 0.08367853611707687
step: 300, loss: 0.06606411188840866
step: 310, loss: 0.2131781429052353
step: 320, loss: 0.05904887616634369
step: 330, loss: 0.04661988839507103
step: 340, loss: 0.4588679075241089
step: 350, loss: 0.15954674780368805
step: 360, loss: 0.04346908628940582
step: 370, loss: 0.08940154314041138
step: 380, loss: 0.3584693670272827
step: 390, loss: 0.06549413502216339
```

```
step: 400, loss: 0.06782669574022293
step: 410, loss: 0.24275223910808563
step: 420, loss: 0.028607342392206192
step: 430, loss: 0.02465187758207321
step: 440, loss: 0.04047304764389992
step: 450, loss: 0.11786191910505295
step: 460, loss: 0.022042684257030487
step: 470, loss: 0.05416087061166763
step: 480, loss: 0.09028230607509613
step: 490, loss: 0.10599192976951599
step: 500, loss: 0.07618474215269089
step: 510, loss: 0.04895693436264992
step: 520, loss: 0.03900996595621109
step: 530, loss: 0.02354273572564125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9272388059701492, f1=0.9260299625468165,
best_f1=0.9260299625468165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08030852675437927
step: 10, loss: 0.08249087631702423
step: 20, loss: 0.06195008382201195
step: 30, loss: 0.010856944136321545
step: 40, loss: 0.12187457829713821
step: 50, loss: 0.12304141372442245
step: 60, loss: 0.012338313274085522
step: 70, loss: 0.025276226922869682
step: 80, loss: 0.04136141389608383
step: 90, loss: 0.06185659021139145
step: 100, loss: 0.10746535658836365
step: 110, loss: 0.008104510605335236
step: 120, loss: 0.029239939525723457
step: 130, loss: 0.02661767415702343
step: 140, loss: 0.13251546025276184
step: 150, loss: 0.05548052117228508
step: 160, loss: 0.06608638912439346
step: 170, loss: 0.09624125808477402
step: 180, loss: 0.06901484727859497
step: 190, loss: 0.02848576195538044
step: 200, loss: 0.2542707622051239
step: 210, loss: 0.043227583169937134
```

```
step: 220, loss: 0.00629703514277935
step: 230, loss: 0.07024337351322174
step: 240, loss: 0.07700271904468536
step: 250, loss: 0.029931366443634033
step: 260, loss: 0.0778263658285141
step: 270, loss: 0.01994609273970127
step: 280, loss: 0.0619329959154129
step: 290, loss: 0.04642203822731972
step: 300, loss: 0.09964308142662048
step: 310, loss: 0.04110448807477951
step: 320, loss: 0.02371993474662304
step: 330, loss: 0.04141673445701599
step: 340, loss: 0.02112097106873989
step: 350, loss: 0.002958935219794512
step: 360, loss: 0.11234230548143387
step: 370, loss: 0.030572835355997086
step: 380, loss: 0.1263577789068222
step: 390, loss: 0.007389975246042013
step: 400, loss: 0.06313920766115189
step: 410, loss: 0.11115492880344391
step: 420, loss: 0.08148985356092453
step: 430, loss: 0.11579343676567078
step: 440, loss: 0.01877456158399582
step: 450, loss: 0.06605809181928635
step: 460, loss: 0.05688653886318207
step: 470, loss: 0.023874858394265175
step: 480, loss: 0.004852067679166794
step: 490, loss: 0.020822888240218163
step: 500, loss: 0.005243777297437191
step: 510, loss: 0.029292207211256027
step: 520, loss: 0.4264015555381775
step: 530, loss: 0.11494992673397064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9386446886446886, f1=0.930939226519337,
best_f1=0.930939226519337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16541731357574463
step: 10, loss: 0.0405702143907547
step: 20, loss: 0.00499356584623456
step: 30, loss: 0.03410978615283966
```

```
step: 40, loss: 0.10071620345115662
step: 50, loss: 0.034317497164011
step: 60, loss: 0.0031827320344746113
step: 70, loss: 0.06338975578546524
step: 80, loss: 0.01668400689959526
step: 90, loss: 0.0915352925658226
step: 100, loss: 0.032366037368774414
step: 110, loss: 0.00586301926523447
step: 120, loss: 0.19387361407279968
step: 130, loss: 0.19017192721366882
step: 140, loss: 0.014800135977566242
step: 150, loss: 0.051931921392679214
step: 160, loss: 0.026817690581083298
step: 170, loss: 0.007864127866923809
step: 180, loss: 0.04505644738674164
step: 190, loss: 0.005729412194341421
step: 200, loss: 0.01289347279816866
step: 210, loss: 0.061309363692998886
step: 220, loss: 0.06605108082294464
step: 230, loss: 0.021137213334441185
step: 240, loss: 0.03100053407251835
step: 250, loss: 0.09094065427780151
step: 260, loss: 0.11393336206674576
step: 270, loss: 0.009295233525335789
step: 280, loss: 0.010028153657913208
step: 290, loss: 0.01634019985795021
step: 300, loss: 0.11920381337404251
step: 310, loss: 0.03173546865582466
step: 320, loss: 0.047911547124385834
step: 330, loss: 0.005088282749056816
step: 340, loss: 0.0029304134659469128
step: 350, loss: 0.042158208787441254
step: 360, loss: 0.016681451350450516
step: 370, loss: 0.045814599841833115
step: 380, loss: 0.024627750739455223
step: 390, loss: 0.02527880296111107
step: 400, loss: 0.11422175914049149
step: 410, loss: 0.017815371975302696
step: 420, loss: 0.008686881512403488
step: 430, loss: 0.020948797464370728
step: 440, loss: 0.20241400599479675
step: 450, loss: 0.045933112502098083
step: 460, loss: 0.07162599265575409
step: 470, loss: 0.013270683586597443
step: 480, loss: 0.048403866589069366
step: 490, loss: 0.015489035286009312
step: 500, loss: 0.008013835176825523
step: 510, loss: 0.025387035682797432
```

```
step: 520, loss: 0.059937190264463425
step: 530, loss: 0.010820649564266205
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9500699953336444, f1=0.9347623485554521,
best_f1=0.9347623485554521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0079558240249753
step: 10, loss: 0.004860066808760166
step: 20, loss: 0.03605053573846817
step: 30, loss: 0.04727673530578613
step: 40, loss: 0.02110331878066063
step: 50, loss: 0.01030502375215292
step: 60, loss: 0.01276878546923399
step: 70, loss: 0.0020476914942264557
step: 80, loss: 0.04476063326001167
step: 90, loss: 0.13394738733768463
step: 100, loss: 0.0036806734278798103
step: 110, loss: 0.04696742817759514
step: 120, loss: 0.00273273722268641
step: 130, loss: 0.15152966976165771
step: 140, loss: 0.048195596784353256
step: 150, loss: 0.005150969140231609
step: 160, loss: 0.0071240635588765144
step: 170, loss: 0.02179255336523056
step: 180, loss: 0.04355870932340622
step: 190, loss: 0.009920582175254822
step: 200, loss: 0.0046072485856711864
step: 210, loss: 0.015483267605304718
step: 220, loss: 0.029412128031253815
step: 230, loss: 0.003076923545449972
step: 240, loss: 0.00273023685404211
step: 250, loss: 0.14533664286136627
step: 260, loss: 0.004651014693081379
step: 270, loss: 0.0339214913547039
step: 280, loss: 0.006874455139040947
step: 290, loss: 0.039286743849515915
step: 300, loss: 0.004580103792250156
step: 310, loss: 0.013379957526922226
step: 320, loss: 0.10375551879405975
step: 330, loss: 0.03458503261208534
```

```
step: 340, loss: 0.022720256820321083
step: 350, loss: 0.17414137721061707
step: 360, loss: 0.008919918909668922
step: 370, loss: 0.017454085871577263
step: 380, loss: 0.022834954783320427
step: 390, loss: 0.0017347523244097829
step: 400, loss: 0.016304470598697662
step: 410, loss: 0.004341879393905401
step: 420, loss: 0.02452838607132435
step: 430, loss: 0.03276090323925018
step: 440, loss: 0.10505934059619904
step: 450, loss: 0.02461911179125309
step: 460, loss: 0.025967685505747795
step: 470, loss: 0.003316279035061598
step: 480, loss: 0.00556403212249279
step: 490, loss: 0.001225331798195839
step: 500, loss: 0.14841145277023315
step: 510, loss: 0.07150329649448395
step: 520, loss: 0.009143432602286339
step: 530, loss: 0.11878790706396103
/usr/local/lib/python3.7/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=0.9488243430152145, f1=0.936111111111111,
best_f1=0.9347623485554521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017577011603862047
step: 10, loss: 0.009328382089734077
step: 20, loss: 0.008558886125683784
step: 30, loss: 0.0031049721874296665
step: 40, loss: 0.006030091550201178
step: 50, loss: 0.031470827758312225
step: 60, loss: 0.017255976796150208
step: 70, loss: 0.0026521612890064716
step: 80, loss: 0.002235390478745103
step: 90, loss: 0.013220120221376419
step: 100, loss: 0.18506084382534027
step: 110, loss: 0.0013725858880206943
step: 120, loss: 0.005351218394935131
step: 130, loss: 0.012544311583042145
step: 140, loss: 0.0009702868992462754
step: 150, loss: 0.016991371288895607
step: 160, loss: 0.021514058113098145
step: 170, loss: 0.06543455272912979
step: 180, loss: 0.013513629324734211
step: 190, loss: 0.0008853071485646069
```

```
step: 200, loss: 0.00256588333286345
step: 210, loss: 0.012309846468269825
step: 220, loss: 0.0002775055472739041
step: 230, loss: 0.0008505308069288731
step: 240, loss: 0.006063788663595915
step: 250, loss: 0.19570611417293549
step: 260, loss: 0.00403072964400053
step: 270, loss: 0.0012605558149516582
step: 280, loss: 0.0048141032457351685
step: 290, loss: 0.0029882341623306274
step: 300, loss: 0.008568324148654938
step: 310, loss: 0.27342602610588074
step: 320, loss: 0.025304753333330154
step: 330, loss: 0.016749251633882523
step: 340, loss: 0.003135389881208539
step: 350, loss: 0.009879658930003643
step: 360, loss: 0.013619767501950264
step: 370, loss: 0.00021048825874458998
step: 380, loss: 0.0007691243081353605
step: 390, loss: 0.0009738680091686547
step: 400, loss: 0.0018704847898334265
step: 410, loss: 0.12956252694129944
step: 420, loss: 0.20943206548690796
step: 430, loss: 0.013086794875562191
step: 440, loss: 0.007383294869214296
step: 450, loss: 0.0035123853012919426
step: 460, loss: 0.005970388185232878
step: 470, loss: 0.0032048409339040518
step: 480, loss: 0.011107800528407097
step: 490, loss: 0.016755232587456703
step: 500, loss: 0.04135247319936752
step: 510, loss: 0.0031773962546139956
step: 520, loss: 0.03541985899209976
step: 530, loss: 0.004172190558165312
/usr/local/lib/python3.7/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=0.9415584415584416, f1=0.9220055710306406,
best_f1=0.9347623485554521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02628444880247116
step: 10, loss: 0.0004543873656075448
step: 20, loss: 0.014605814591050148
step: 30, loss: 0.0013436758890748024
step: 40, loss: 0.0005964335286989808
step: 50, loss: 0.00035473430762067437
```

```
step: 60, loss: 0.0015356147196143866
step: 70, loss: 0.00011032048496417701
step: 80, loss: 0.00010539736831560731
step: 90, loss: 0.0046731941401958466
step: 100, loss: 0.035956982523202896
step: 110, loss: 0.005140622612088919
step: 120, loss: 0.004533456172794104
step: 130, loss: 0.004518011584877968
step: 140, loss: 0.00036457492387853563
step: 150, loss: 0.00062024517683123
step: 160, loss: 0.0829303041100502
step: 170, loss: 0.006902300752699375
step: 180, loss: 0.001346240402199328
step: 190, loss: 0.18361791968345642
step: 200, loss: 0.06986705213785172
step: 210, loss: 0.001617090543732047
step: 220, loss: 0.006792394444346428
step: 230, loss: 0.005773358047008514
step: 240, loss: 0.0006590512930415571
step: 250, loss: 0.009208919480443
step: 260, loss: 0.001700229593552649
step: 270, loss: 0.002130756387487054
step: 280, loss: 0.008478320203721523
step: 290, loss: 0.11228358745574951
step: 300, loss: 0.02574084885418415
step: 310, loss: 0.010411299765110016
step: 320, loss: 5.00918977195397e-05
step: 330, loss: 0.00281047192402184
step: 340, loss: 0.0005805540131404996
step: 350, loss: 0.0005619022413156927
step: 360, loss: 0.2100093811750412
step: 370, loss: 0.017859697341918945
step: 380, loss: 0.006485898979008198
step: 390, loss: 0.00854035746306181
step: 400, loss: 0.01427190937101841
step: 410, loss: 0.001489780261181295
step: 420, loss: 0.004361957311630249
step: 430, loss: 0.000463138974737376
step: 440, loss: 0.010102776810526848
step: 450, loss: 0.17999932169914246
step: 460, loss: 0.004768471699208021
step: 470, loss: 0.005796907003968954
step: 480, loss: 0.006487810052931309
step: 490, loss: 0.010192757472395897
step: 500, loss: 0.004545341245830059
step: 510, loss: 0.004701352212578058
step: 520, loss: 0.003415603656321764
step: 530, loss: 0.0006999254110269248
```

```
/usr/local/lib/python3.7/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=0.9455909943714823, f1=0.9363295880149813,
best_f1=0.9347623485554521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000911843788344413
step: 10, loss: 0.04870578274130821
step: 20, loss: 0.0026529559399932623
step: 30, loss: 0.005327289924025536
step: 40, loss: 0.0015346321742981672
step: 50, loss: 0.032856591045856476
step: 60, loss: 0.0016612776089459658
step: 70, loss: 0.0004856786981690675
step: 80, loss: 0.0003790585324168205
step: 90, loss: 0.00014168652705848217
step: 100, loss: 0.003539229044690728
step: 110, loss: 0.0042801653034985065
step: 120, loss: 0.000895923760253936
step: 130, loss: 0.00019533885642886162
step: 140, loss: 0.00019298984261695296
step: 150, loss: 0.0010469552362337708
step: 160, loss: 0.0047536008059978485
step: 170, loss: 0.005168461240828037
step: 180, loss: 0.06988606601953506
step: 190, loss: 0.011578306555747986
step: 200, loss: 0.0001824118080548942
step: 210, loss: 0.0022756345570087433
step: 220, loss: 0.0002808604622259736
step: 230, loss: 0.0013956796610727906
step: 240, loss: 0.020630592480301857
step: 250, loss: 0.001884455094113946
step: 260, loss: 0.0041137379594147205
step: 270, loss: 0.00022699924011249095
step: 280, loss: 0.0010043744696304202
step: 290, loss: 0.01478970143944025
step: 300, loss: 0.0005453634657897055
step: 310, loss: 0.0015641215723007917
step: 320, loss: 0.0015946927014738321
step: 330, loss: 0.021016914397478104
step: 340, loss: 0.0002540763234719634
step: 350, loss: 0.004248789045959711
step: 360, loss: 0.05549727380275726
step: 370, loss: 0.00518184807151556
step: 380, loss: 0.006209513638168573
step: 390, loss: 6.585507071577013e-05
```

```
step: 400, loss: 0.010691892355680466
step: 410, loss: 0.027321552857756615
step: 420, loss: 0.005951267667114735
step: 430, loss: 0.0013451435370370746
step: 440, loss: 0.03579248860478401
step: 450, loss: 0.08485225588083267
step: 460, loss: 0.004050090909004211
step: 470, loss: 0.09811216592788696
step: 480, loss: 0.01331216562539339
step: 490, loss: 0.0007997087086550891
step: 500, loss: 0.04475023224949837
step: 510, loss: 0.0008782123913988471
step: 520, loss: 0.0014122247230261564
step: 530, loss: 0.00544670270755887
/usr/local/lib/python3.7/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=0.949767441860465, f1=0.9411764705882353,
best_f1=0.9347623485554521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025820608716458082
step: 10, loss: 0.0005635181441903114
step: 20, loss: 3.300380194559693e-05
step: 30, loss: 0.0014345480594784021
step: 40, loss: 3.314968853374012e-05
step: 50, loss: 0.0010893942089751363
step: 60, loss: 0.0049570598639547825
step: 70, loss: 0.00012886202603112906
step: 80, loss: 0.0558524988591671
step: 90, loss: 9.310048335464671e-05
step: 100, loss: 0.0007630935870110989
step: 110, loss: 0.0005313607398420572
step: 120, loss: 0.0007726183393970132
step: 130, loss: 0.00016723405860830098
step: 140, loss: 3.7871453969273716e-05
step: 150, loss: 0.0004104974796064198
step: 160, loss: 5.973411680315621e-05
step: 170, loss: 0.02035900391638279
step: 180, loss: 6.787067104596645e-05
step: 190, loss: 0.04007435217499733
step: 200, loss: 0.07696213573217392
step: 210, loss: 0.05943530425429344
step: 220, loss: 0.001688160584308207
step: 230, loss: 0.05462021380662918
step: 240, loss: 0.001939344103448093
step: 250, loss: 0.006743705365806818
```

```
step: 260, loss: 2.2656618966720998e-05
step: 270, loss: 0.011391456238925457
step: 280, loss: 3.715581988217309e-05
step: 290, loss: 0.00015246937982738018
step: 300, loss: 3.059363370994106e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 310, loss: 2.9096996513544582e-05
step: 320, loss: 2.4455592210870236e-05
step: 330, loss: 3.224085958208889e-05
step: 340, loss: 0.009915957227349281
step: 350, loss: 2.9271599487401545e-05
step: 360, loss: 0.04566660523414612
step: 370, loss: 0.03895631805062294
step: 380, loss: 2.3945136490510777e-05
step: 390, loss: 0.004640047438442707
step: 400, loss: 0.00716444430872798
step: 410, loss: 5.6824494095053524e-05
step: 420, loss: 8.284737850772217e-05
step: 430, loss: 4.385627471492626e-05
step: 440, loss: 0.0014745097141712904
step: 450, loss: 0.05705399811267853
step: 460, loss: 0.008640236221253872
step: 470, loss: 0.027749596163630486
step: 480, loss: 0.0035412581637501717
step: 490, loss: 0.014656739309430122
step: 500, loss: 0.0013228283496573567
step: 510, loss: 0.00047657277900725603
step: 520, loss: 0.009406558237969875
step: 530, loss: 4.255831299815327e-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=0.9468331021729081, f1=0.9377018920166128,
best_f1=0.9347623485554521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.901662108139135e-05
step: 10, loss: 0.005919867195188999
step: 20, loss: 2.737218892434612e-05
step: 30, loss: 0.04944134131073952
step: 40, loss: 6.974772259127349e-05
step: 50, loss: 0.0002561717410571873
step: 60, loss: 7.888447726145387e-05
step: 70, loss: 2.3297248844755813e-05
step: 80, loss: 0.036424264311790466
step: 90, loss: 0.05346068739891052
step: 100, loss: 0.010598599910736084
```

```
step: 110, loss: 0.006018141284584999
step: 120, loss: 2.4824594220262952e-05
step: 130, loss: 3.976110747316852e-05
step: 140, loss: 3.177571852575056e-05
step: 150, loss: 0.013722908683121204
step: 160, loss: 0.0011619686847552657
step: 170, loss: 0.009089814499020576
step: 180, loss: 0.008441564626991749
step: 190, loss: 4.0329614421352744e-05
step: 200, loss: 0.004623837769031525
step: 210, loss: 1.8071035810862668e-05
step: 220, loss: 0.000683385063894093
step: 230, loss: 0.0045383949764072895
step: 240, loss: 6.680592196062207e-05
step: 250, loss: 6.319354724837467e-05
step: 260, loss: 0.00011180284491274506
step: 270, loss: 1.956853157025762e-05
step: 280, loss: 0.001205754466354847
step: 290, loss: 2.292849967489019e-05
step: 300, loss: 1.8868173356167972e-05
step: 310, loss: 0.0001267590414499864
step: 320, loss: 2.268232492497191e-05
step: 330, loss: 0.0005039031966589391
step: 340, loss: 0.0007438199245370924
step: 350, loss: 0.043415676802396774
step: 360, loss: 0.002475918270647526
step: 370, loss: 0.00036292799632065
step: 380, loss: 0.000788328645285219
step: 390, loss: 1.5601232007611543e-05
step: 400, loss: 0.014172581024467945
step: 410, loss: 0.000221638212678954
step: 420, loss: 0.006626340560615063
step: 430, loss: 0.0225802194327116
step: 440, loss: 2.759848030109424e-05
step: 450, loss: 0.0017540967091917992
step: 460, loss: 0.0027396243531256914
step: 470, loss: 4.6962039050413296e-05
step: 480, loss: 0.000131715860334225
step: 490, loss: 0.004170346539467573
step: 500, loss: 0.00019484020594973117
step: 510, loss: 0.001912153558805585
step: 520, loss: 0.00197864742949605
step: 530, loss: 0.2811165452003479
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9526932084309133, f1=0.9382022471910113,
best_f1=0.9382022471910113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005741350469179451
step: 10, loss: 0.0015470301732420921
step: 20, loss: 0.0005520563572645187
step: 30, loss: 0.0009257713568201822
step: 40, loss: 0.00011275332508375868
step: 50, loss: 0.0032566317822784185
step: 60, loss: 0.02827211283147335
step: 70, loss: 0.0003762084525078535
step: 80, loss: 0.00027338971267454326
step: 90, loss: 0.00023090923787094653
step: 100, loss: 0.0005474633071571589
step: 110, loss: 0.0018051525112241507
step: 120, loss: 5.9035584854427725e-05
step: 130, loss: 0.0007092729210853577
step: 140, loss: 0.00016092017176561058
step: 150, loss: 0.0002988879568874836
step: 160, loss: 0.04491647332906723
step: 170, loss: 0.009691908955574036
step: 180, loss: 0.0007120638038031757
step: 190, loss: 4.2506380850682035e-05
step: 200, loss: 0.009529788047075272
step: 210, loss: 0.00550646148622036
step: 220, loss: 0.0010808203369379044
step: 230, loss: 6.284754635998979e-05
step: 240, loss: 0.00013323577877599746
step: 250, loss: 0.019243741407990456
step: 260, loss: 0.006032528821378946
step: 270, loss: 0.00010775921691674739
step: 280, loss: 0.038335997611284256
step: 290, loss: 0.019770296290516853
step: 300, loss: 3.9329312130576e-05
step: 310, loss: 0.025741633027791977
step: 320, loss: 0.001790503622032702
step: 330, loss: 0.0003572499263100326
step: 340, loss: 1.542975769552868e-05
step: 350, loss: 0.0006071269745007157
step: 360, loss: 0.00014077678497415036
step: 370, loss: 3.2466563425259665e-05
step: 380, loss: 0.0006168561521917582
step: 390, loss: 2.4016082534217276e-05
step: 400, loss: 5.125419556861743e-05
```

```
step: 410, loss: 0.00010059457417810336
step: 420, loss: 2.1628342437907122e-05
step: 430, loss: 1.664045157667715e-05
step: 440, loss: 2.3319316824199632e-05
step: 450, loss: 0.05508232116699219
step: 460, loss: 2.2764375898987055e-05
step: 470, loss: 0.07546018064022064
step: 480, loss: 5.2552641136571765e-05
step: 490, loss: 3.163137807860039e-05
step: 500, loss: 6.631668657064438e-05
step: 510, loss: 3.772874333662912e-05
step: 520, loss: 0.014758149161934853
step: 530, loss: 0.020134154707193375
/usr/local/lib/python3.7/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=0.9503480278422275, f1=0.9344413665743306,
best_f1=0.9382022471910113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2380807422450744e-05
step: 10, loss: 2.478689384588506e-05
step: 20, loss: 0.001030799001455307
step: 30, loss: 2.9551216357504018e-05
step: 40, loss: 2.3184624296845868e-05
step: 50, loss: 2.3460392185370438e-05
step: 60, loss: 2.860975837393198e-05
step: 70, loss: 3.2363990612793714e-05
step: 80, loss: 6.063817272661254e-05
step: 90, loss: 2.6264266125508584e-05
step: 100, loss: 1.8944996554637328e-05
step: 110, loss: 3.0235949452617206e-05
step: 120, loss: 2.9149072361178696e-05
step: 130, loss: 1.5161615010583773e-05
step: 140, loss: 1.6305226381518878e-05
step: 150, loss: 6.920178566360846e-05
step: 160, loss: 0.00025441887555643916
step: 170, loss: 1.0546143130341079e-05
step: 180, loss: 1.5232371879392304e-05
step: 190, loss: 3.2644817110849544e-05
step: 200, loss: 1.1406622434151359e-05
step: 210, loss: 1.371615599055076e-05
step: 220, loss: 0.002951847855001688
step: 230, loss: 1.677351065154653e-05
step: 240, loss: 0.0030061746947467327
step: 250, loss: 0.11444848775863647
step: 260, loss: 0.0035571043845266104
```

```
step: 270, loss: 0.0035938250366598368
step: 280, loss: 0.0007452171994373202
step: 290, loss: 0.00013497876352630556
step: 300, loss: 0.00018484951579011977
step: 310, loss: 0.001703393878415227
step: 320, loss: 0.0008201346499845386
step: 330, loss: 0.00034729568869806826
step: 340, loss: 0.00293814018368721
step: 350, loss: 0.08095315098762512
step: 360, loss: 0.0008466150611639023
step: 370, loss: 0.0007597003714181483
step: 380, loss: 0.001243028906174004
step: 390, loss: 0.0007575678173452616
step: 400, loss: 0.010966328904032707
step: 410, loss: 0.0015017892001196742
step: 420, loss: 3.1083291105460376e-05
step: 430, loss: 0.01629013940691948
step: 440, loss: 3.0785253329668194e-05
step: 450, loss: 3.668148201541044e-05
step: 460, loss: 3.7478821468539536e-05
step: 470, loss: 7.741766603430733e-05
step: 480, loss: 0.0015107013750821352
step: 490, loss: 2.584937828942202e-05
step: 500, loss: 2.760387542366516e-05
step: 510, loss: 0.0008590848883613944
step: 520, loss: 5.7784196542343125e-05
step: 530, loss: 0.007409730460494757
/usr/local/lib/python3.7/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=0.9498389323515877, f1=0.9327231121281464,
best_f1=0.9382022471910113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9279770387802273e-05
step: 10, loss: 0.002015364822000265
step: 20, loss: 0.0008172008092515171
step: 30, loss: 2.08725432457868e-05
step: 40, loss: 9.63663260336034e-05
step: 50, loss: 0.0004884536028839648
step: 60, loss: 3.998198008048348e-05
step: 70, loss: 0.00130382262259215117
step: 80, loss: 4.765037738252431e-05
step: 90, loss: 0.0014241632306948304
step: 100, loss: 0.00463535450398922
step: 110, loss: 0.0009640150237828493
step: 120, loss: 0.0001365148345939815
```

```
step: 130, loss: 0.00041697744745761156
step: 140, loss: 3.6903507862007245e-05
step: 150, loss: 0.008573058992624283
step: 160, loss: 0.0006740653188899159
step: 170, loss: 2.4180309992516413e-05
step: 180, loss: 3.7871144741075113e-05
step: 190, loss: 4.077446646988392e-05
step: 200, loss: 0.0006973338895477355
step: 210, loss: 2.3781798518029973e-05
step: 220, loss: 2.4437053070869297e-05
step: 230, loss: 2.793440580717288e-05
step: 240, loss: 0.00017798769113142043
step: 250, loss: 0.00014922766422387213
step: 260, loss: 2.762529948086012e-05
step: 270, loss: 0.0012932714307680726
step: 280, loss: 6.125099025666714e-05
step: 290, loss: 0.00042477832175791264
step: 300, loss: 3.388883123989217e-05
step: 310, loss: 5.764975503552705e-05
step: 320, loss: 0.0001824084756663069
step: 330, loss: 0.0011556919198483229
step: 340, loss: 1.6078132830443792e-05
step: 350, loss: 1.8111693862010725e-05
step: 360, loss: 2.9248734790598974e-05
step: 370, loss: 1.7169433704111725e-05
step: 380, loss: 4.582684050546959e-05
step: 390, loss: 5.0236936658620834e-05
step: 400, loss: 4.714592796517536e-05
step: 410, loss: 2.353574564040173e-05
step: 420, loss: 1.9244536815676838e-05
step: 430, loss: 1.8048795027425513e-05
step: 440, loss: 0.000594829733017832
step: 450, loss: 0.0006910130032338202
step: 460, loss: 0.00020278881129343063
step: 470, loss: 0.017298104241490364
step: 480, loss: 1.933025305334013e-05
step: 490, loss: 1.830209293984808e-05
step: 500, loss: 1.8998624000232667e-05
step: 510, loss: 2.857494109775871e-05
step: 520, loss: 0.010191911831498146
step: 530, loss: 0.010074837133288383
/usr/local/lib/python3.7/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=0.9496738117427772, f1=0.9329608938547487,
best_f1=0.9382022471910113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.5616249584127218e-05
step: 10, loss: 2.1278216081555e-05
step: 20, loss: 0.0006651837611570954
step: 30, loss: 1.6033431165851653e-05
step: 40, loss: 1.6622036127955653e-05
step: 50, loss: 0.0002789173449855298
step: 60, loss: 1.7124833902926184e-05
step: 70, loss: 1.8029893908533268e-05
step: 80, loss: 5.982847505947575e-05
step: 90, loss: 3.468488284852356e-05
step: 100, loss: 0.0005791387520730495
step: 110, loss: 1.2852141480834689e-05
step: 120, loss: 0.00011348746193107218
step: 130, loss: 1.142910241469508e-05
step: 140, loss: 1.892036925710272e-05
step: 150, loss: 2.8361810109345242e-05
step: 160, loss: 0.02049066126346588
step: 170, loss: 0.00018485482723917812
step: 180, loss: 0.0005864240229129791
step: 190, loss: 1.4621607078879606e-05
step: 200, loss: 2.5829429432633333e-05
step: 210, loss: 0.000104366714367643
step: 220, loss: 2.054045398836024e-05
step: 230, loss: 0.002698298543691635
step: 240, loss: 1.710251672193408e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 250, loss: 0.000751926563680172
step: 260, loss: 1.6107955161714926e-05
step: 270, loss: 0.00019664419232867658
step: 280, loss: 1.6275544112431817e-05
step: 290, loss: 1.1555761375348084e-05
step: 300, loss: 2.249204226245638e-05
step: 310, loss: 1.4375707905855961e-05
step: 320, loss: 5.2135303121758625e-05
step: 330, loss: 5.624391997116618e-05
step: 340, loss: 1.499035170127172e-05
step: 350, loss: 9.99861367745325e-06
step: 360, loss: 0.007497929502278566
step: 370, loss: 1.0315255167370196e-05
step: 380, loss: 1.6115214748424478e-05
step: 390, loss: 1.2230019819980953e-05
step: 400, loss: 1.4308666322904173e-05
step: 410, loss: 5.1778271881630644e-05
step: 420, loss: 1.1812786397058517e-05
step: 430, loss: 1.786979555618018e-05
step: 440, loss: 1.2092172255506739e-05
step: 450, loss: 1.5742734831292182e-05
```

```
step: 460, loss: 0.014653470367193222
step: 470, loss: 4.160860407864675e-05
step: 480, loss: 9.491970558883622e-06
step: 490, loss: 1.220391186507186e-05
step: 500, loss: 1.250553032150492e-05
step: 510, loss: 9.979982678487431e-06
step: 520, loss: 0.0004578513326123357
step: 530, loss: 1.1790419193857815e-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=0.9495136637332099, f1=0.9364269141531322,
best_f1=0.9382022471910113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.8479384127422236e-05
step: 10, loss: 1.0751089575933293e-05
step: 20, loss: 1.1607878150243778e-05
step: 30, loss: 0.10138686001300812
step: 40, loss: 1.2002760740870144e-05
step: 50, loss: 0.00010387397196609527
step: 60, loss: 4.360793900559656e-05
step: 70, loss: 1.485614848206751e-05
step: 80, loss: 3.452270175330341e-05
step: 90, loss: 9.871942893369123e-06
step: 100, loss: 7.700575952185318e-05
step: 110, loss: 0.010248151607811451
step: 120, loss: 1.5534129488514736e-05
step: 130, loss: 1.1429096048232168e-05
step: 140, loss: 1.6990619769785553e-05
step: 150, loss: 1.2375304322631564e-05
step: 160, loss: 2.246166877739597e-05
step: 170, loss: 0.00010447037493577227
step: 180, loss: 1.1663729310384952e-05
step: 190, loss: 0.0006062902975827456
step: 200, loss: 1.2323144801484887e-05
step: 210, loss: 0.011265791952610016
step: 220, loss: 9.503150977252517e-06
step: 230, loss: 1.1358303709130269e-05
step: 240, loss: 0.0011457817163318396
step: 250, loss: 2.4945305995061062e-05
step: 260, loss: 0.00014669295342173427
step: 270, loss: 1.5970015738275833e-05
step: 280, loss: 3.525482316035777e-05
step: 290, loss: 0.00028258824022486806
step: 300, loss: 3.315725189168006e-05
step: 310, loss: 2.191378371207975e-05
```

```
step: 320, loss: 5.080296614323743e-05
step: 330, loss: 0.001647485070861876
step: 340, loss: 0.00044679385609924793
step: 350, loss: 1.0277995897922665e-05
step: 360, loss: 1.1991551218670793e-05
step: 370, loss: 4.9665875849314034e-05
step: 380, loss: 2.2600512238568626e-05
step: 390, loss: 8.206767233787104e-06
step: 400, loss: 0.003454432124271989
step: 410, loss: 1.24607358884532e-05
step: 420, loss: 1.1082606761192437e-05
step: 430, loss: 0.01095820963382721
step: 440, loss: 1.678900844126474e-05
step: 450, loss: 1.0892640602833126e-05
step: 460, loss: 0.02123197354376316
step: 470, loss: 1.3816907994623762e-05
step: 480, loss: 1.0080554602609482e-05
step: 490, loss: 1.1809065654233564e-05
step: 500, loss: 0.00010848757665371522
step: 510, loss: 0.008187178522348404
step: 520, loss: 8.385575711145066e-06
step: 530, loss: 1.244978830072796e-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=0.9512419503219871, f1=0.9315571887919155,
best_f1=0.9382022471910113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.214055282616755e-05
step: 10, loss: 1.7783888324629515e-05
step: 20, loss: 0.0034210151061415672
step: 30, loss: 0.0008808679413050413
step: 40, loss: 1.2185279047116637e-05
step: 50, loss: 1.9739249182748608e-05
step: 60, loss: 2.9240425647003576e-05
step: 70, loss: 1.027795769914519e-05
step: 80, loss: 9.614897862775251e-06
step: 90, loss: 9.890560249914415e-06
step: 100, loss: 7.420739166263957e-06
step: 110, loss: 1.1414183063607197e-05
step: 120, loss: 1.0151328751817346e-05
step: 130, loss: 2.380518344580196e-05
step: 140, loss: 0.00017718601156957448
step: 150, loss: 1.2181533747934736e-05
step: 160, loss: 0.009671803563833237
step: 170, loss: 7.700129572185688e-06
```

```
step: 180, loss: 1.547449573990889e-05
step: 190, loss: 0.08614730089902878
step: 200, loss: 0.0006649502320215106
step: 210, loss: 1.1481087312859017e-05
step: 220, loss: 1.0676588317437563e-05
step: 230, loss: 1.3384811609284952e-05
step: 240, loss: 9.626076462154742e-06
step: 250, loss: 9.52174832491437e-06
step: 260, loss: 9.667041013017297e-06
step: 270, loss: 1.1097527021775022e-05
step: 280, loss: 1.2233654160809238e-05
step: 290, loss: 1.419273394276388e-05
step: 300, loss: 1.809962122933939e-05
step: 310, loss: 0.0005782514926977456
step: 320, loss: 1.3224563190306071e-05
step: 330, loss: 1.3965825928607956e-05
step: 340, loss: 1.2665836038650014e-05
step: 350, loss: 0.012308921664953232
step: 360, loss: 1.0710122296586633e-05
step: 370, loss: 0.00011016183270839974
step: 380, loss: 1.0460526027600281e-05
step: 390, loss: 9.752734513313044e-06
step: 400, loss: 1.2490759218053427e-05
step: 410, loss: 1.268060805159621e-05
step: 420, loss: 1.006192178465426e-05
step: 430, loss: 7.703858500462957e-06
step: 440, loss: 1.0769701475510374e-05
step: 450, loss: 0.01009879820048809
step: 460, loss: 2.6248610083712265e-05
step: 470, loss: 9.767628398549277e-06
step: 480, loss: 0.007797719445079565
step: 490, loss: 9.808609320316464e-06
step: 500, loss: 0.00018945841293316334
step: 510, loss: 8.571841135562863e-06
step: 520, loss: 1.0937351362372283e-05
step: 530, loss: 8.791631444182713e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9499536607970344, f1=0.935064935064935,
best_f1=0.9382022471910113
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:21, 266.82it/s]
load_f1 = 0.9520295202952029
real_f1 = 0.9494485294117647
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 207.54it/s]
```

3.1.4 Amazon-Google - Running the matcher

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

```
  --save_model \
  --dk general  \
  --dk product \
  --summarize


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

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

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

Defaults for this optimization level are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5595887303352356
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.3648218810558319
step: 20, loss: 0.439184308052063
step: 30, loss: 0.40096181631088257
step: 40, loss: 0.32461291551589966
step: 50, loss: 0.419864296913147
step: 60, loss: 0.49358534812927246
step: 70, loss: 0.3032682240009308
step: 80, loss: 0.3611461818218231
step: 90, loss: 0.22359588742256165
step: 100, loss: 0.2473190873861313
step: 110, loss: 0.30829304456710815
step: 120, loss: 0.3833301067352295
step: 130, loss: 0.2878572642803192
step: 140, loss: 0.4590226709842682
step: 150, loss: 0.32024145126342773
step: 160, loss: 0.5575697422027588
step: 170, loss: 0.21979054808616638
step: 180, loss: 0.29633983969688416
step: 190, loss: 0.5842649936676025
step: 200, loss: 0.24490970373153687
step: 210, loss: 0.3031575381755829
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4110787172011661, f1=0.4444444444444444,
best_f1=0.4444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.298577219247818
step: 10, loss: 0.08557453751564026
step: 20, loss: 0.44851166009902954
step: 30, loss: 0.41481295228004456
step: 40, loss: 0.39029765129089355
step: 50, loss: 0.2207290530204773
step: 60, loss: 0.5078365206718445
step: 70, loss: 0.3387583792209625
step: 80, loss: 0.2142334282398224
step: 90, loss: 0.15699756145477295
step: 100, loss: 0.5128507614135742
step: 110, loss: 0.30961665511131287
step: 120, loss: 0.13867899775505066
step: 130, loss: 0.15839827060699463
```

```
step: 140, loss: 0.16667567193508148
step: 150, loss: 0.2785812020301819
step: 160, loss: 0.10497947782278061
step: 170, loss: 0.3041667938232422
step: 180, loss: 0.20335876941680908
step: 190, loss: 0.3257316052913666
step: 200, loss: 0.12264944612979889
step: 210, loss: 0.18637390434741974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.530701754385965, f1=0.49756097560975615,
best_f1=0.49756097560975615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06347097456455231
step: 10, loss: 0.1751544028520584
step: 20, loss: 0.3272581398487091
step: 30, loss: 0.09418075531721115
step: 40, loss: 0.40534162521362305
step: 50, loss: 0.24894481897354126
step: 60, loss: 0.24724292755126953
step: 70, loss: 0.12020999938249588
step: 80, loss: 0.23074746131896973
step: 90, loss: 0.10343632847070694
step: 100, loss: 0.3462736904621124
step: 110, loss: 0.0485033243894577
step: 120, loss: 0.1414560228586197
step: 130, loss: 0.17320676147937775
step: 140, loss: 0.1499023586511612
step: 150, loss: 0.19328057765960693
step: 160, loss: 0.08728020638227463
step: 170, loss: 0.15293709933757782
step: 180, loss: 0.10970566421747208
step: 190, loss: 0.050679389387369156
step: 200, loss: 0.08637773245573044
step: 210, loss: 0.2846240699291229
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6045627376425856, f1=0.575875486381323,
best_f1=0.575875486381323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08546736091375351
step: 10, loss: 0.08059665560722351
step: 20, loss: 0.03615427017211914
step: 30, loss: 0.10495668649673462
step: 40, loss: 0.05784149840474129
step: 50, loss: 0.12529224157333374
step: 60, loss: 0.13143882155418396
step: 70, loss: 0.07272949069738388
step: 80, loss: 0.08528143167495728
step: 90, loss: 0.11193327605724335
step: 100, loss: 0.2403509020805359
step: 110, loss: 0.18209624290466309
step: 120, loss: 0.11945502460002899
step: 130, loss: 0.1684512346982956
step: 140, loss: 0.3212917149066925
step: 150, loss: 0.08893769234418869
step: 160, loss: 0.26589569449424744
step: 170, loss: 0.09321105480194092
step: 180, loss: 0.07584012299776077
step: 190, loss: 0.11561375856399536
step: 200, loss: 0.0517815463244915
step: 210, loss: 0.2732938230037689
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6162162162162163, f1=0.6178217821782177,
best_f1=0.6178217821782177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10358358919620514
step: 10, loss: 0.03499981015920639
step: 20, loss: 0.3843635618686676
step: 30, loss: 0.04184955358505249
step: 40, loss: 0.0743248239159584
step: 50, loss: 0.29053449630737305
step: 60, loss: 0.09104415774345398
step: 70, loss: 0.13615885376930237
```

```
step: 80, loss: 0.14694175124168396
step: 90, loss: 0.031039729714393616
step: 100, loss: 0.04564807191491127
step: 110, loss: 0.10865803062915802
step: 120, loss: 0.07670285552740097
step: 130, loss: 0.12649093568325043
step: 140, loss: 0.29593098163604736
step: 150, loss: 0.05603436380624771
step: 160, loss: 0.11292216181755066
step: 170, loss: 0.07321851700544357
step: 180, loss: 0.14667798578739166
step: 190, loss: 0.11137744784355164
step: 200, loss: 0.12288375198841095
step: 210, loss: 0.10275815427303314
/usr/local/lib/python3.7/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=0.595667870036101, f1=0.5875706214689265,
best_f1=0.6178217821782177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04277554899454117
step: 10, loss: 0.04183997958898544
step: 20, loss: 0.012986302375793457
step: 30, loss: 0.2143399566411972
step: 40, loss: 0.07196030020713806
step: 50, loss: 0.11355403810739517
step: 60, loss: 0.07585977762937546
step: 70, loss: 0.13625121116638184
step: 80, loss: 0.05209362506866455
step: 90, loss: 0.07470081001520157
step: 100, loss: 0.1804240643978119
step: 110, loss: 0.030588209629058838
step: 120, loss: 0.030919287353754044
step: 130, loss: 0.026264525949954987
step: 140, loss: 0.19109562039375305
step: 150, loss: 0.07804709672927856
step: 160, loss: 0.06645241379737854
step: 170, loss: 0.06742902845144272
step: 180, loss: 0.0794694796204567
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.033065468072891235
step: 200, loss: 0.1747356802225113
step: 210, loss: 0.10453008860349655
/usr/local/lib/python3.7/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=0.6117647058823529, f1=0.5995717344753747,
best_f1=0.6178217821782177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07707519829273224
step: 10, loss: 0.006104599684476852
step: 20, loss: 0.010320169851183891
step: 30, loss: 0.03782355412840843
step: 40, loss: 0.032190244644880295
step: 50, loss: 0.08824409544467926
step: 60, loss: 0.021830245852470398
step: 70, loss: 0.03785136714577675
step: 80, loss: 0.050781477242708206
step: 90, loss: 0.015070079825818539
step: 100, loss: 0.07489731162786484
step: 110, loss: 0.062012214213609695
step: 120, loss: 0.08655345439910889
step: 130, loss: 0.40928736329078674
step: 140, loss: 0.09988309442996979
step: 150, loss: 0.06286130845546722
step: 160, loss: 0.09811471402645111
step: 170, loss: 0.009096933528780937
step: 180, loss: 0.0032714873086611
step: 190, loss: 0.1405072659254074
step: 200, loss: 0.04238863289356232
step: 210, loss: 0.20566387474536896
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6224899598393574, f1=0.5877192982456141,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09152769297361374
step: 10, loss: 0.01580190286040306
step: 20, loss: 0.015388250350952148
step: 30, loss: 0.06737042218446732
step: 40, loss: 0.10132197290658951
step: 50, loss: 0.0024191278498619795
step: 60, loss: 0.0019737184047698975
step: 70, loss: 0.16744554042816162
step: 80, loss: 0.0643060952425003
```

```
step: 90, loss: 0.05137718468904495
step: 100, loss: 0.0762142762541771
step: 110, loss: 0.021129518747329712
step: 120, loss: 0.13773389160633087
step: 130, loss: 0.002045177621766925
step: 140, loss: 0.02844196744263172
step: 150, loss: 0.108134426176548
step: 160, loss: 0.24243666231632233
step: 170, loss: 0.04880763590335846
step: 180, loss: 0.04962040111422539
step: 190, loss: 0.010529703460633755
step: 200, loss: 0.055925291031599045
step: 210, loss: 0.02449231781065464
/usr/local/lib/python3.7/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=0.6042884990253411, f1=0.6159844054580896,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0750129371881485
step: 10, loss: 0.15912117063999176
step: 20, loss: 0.04528019204735756
step: 30, loss: 0.007835470139980316
step: 40, loss: 0.023944510146975517
step: 50, loss: 0.014847826212644577
step: 60, loss: 0.0007442609639838338
step: 70, loss: 0.0635644719004631
step: 80, loss: 0.13647852838039398
step: 90, loss: 0.02593180164694786
step: 100, loss: 0.028996549546718597
step: 110, loss: 0.0632871761918068
step: 120, loss: 0.10247473418712616
step: 130, loss: 0.04743608087301254
step: 140, loss: 0.05972554534673691
step: 150, loss: 0.03842443227767944
step: 160, loss: 0.03939465060830116
step: 170, loss: 0.028607215732336044
step: 180, loss: 0.031404223293066025
step: 190, loss: 0.0020451906602829695
step: 200, loss: 0.017577778548002243
step: 210, loss: 0.021900251507759094
/usr/local/lib/python3.7/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=0.618042226487524, f1=0.6344294003868473,
best_f1=0.5877192982456141
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015018545091152191
step: 10, loss: 0.013198732398450375
step: 20, loss: 0.008579071611166
step: 30, loss: 0.0024351542815566063
step: 40, loss: 0.0021730619482696056
step: 50, loss: 0.058970604091882706
step: 60, loss: 0.0024800864048302174
step: 70, loss: 0.021260689944028854
step: 80, loss: 0.010457785800099373
step: 90, loss: 0.023801028728485107
step: 100, loss: 0.022464800626039505
step: 110, loss: 0.00233479798771441
step: 120, loss: 0.18233893811702728
step: 130, loss: 0.08649972081184387
step: 140, loss: 0.00385707407258451
step: 150, loss: 0.024975523352622986
step: 160, loss: 0.004411628004163504
step: 170, loss: 0.009169594384729862
step: 180, loss: 0.02931276150047779
step: 190, loss: 0.0032100984826684
step: 200, loss: 0.02294132672250271
step: 210, loss: 0.006592637859284878
/usr/local/lib/python3.7/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=0.6106346483704975, f1=0.6284722222222222,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04494178295135498
step: 10, loss: 0.0114709148183465
step: 20, loss: 0.002880935790017247
step: 30, loss: 0.021219976246356964
step: 40, loss: 0.09204275161027908
step: 50, loss: 0.05178043618798256
step: 60, loss: 0.013400556519627571
step: 70, loss: 0.012295767664909363
step: 80, loss: 0.049491800367832184
step: 90, loss: 0.20257540047168732
step: 100, loss: 0.018757540732622147
step: 110, loss: 0.022621596232056618
step: 120, loss: 0.012543109245598316
step: 130, loss: 0.05853104963898659
step: 140, loss: 0.1265171617269516
```

```
step: 150, loss: 0.0016992138698609709
step: 160, loss: 0.0011493988567963243
step: 170, loss: 0.027485983446240425
step: 180, loss: 0.0017369250999763608
step: 190, loss: 0.10855662077665329
step: 200, loss: 0.007705645635724068
step: 210, loss: 0.05103325471282005
/usr/local/lib/python3.7/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=0.5978947368421053, f1=0.616052060737527,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002316004829481244
step: 10, loss: 0.03301035240292549
step: 20, loss: 0.0031466712243855
step: 30, loss: 0.00035161522100679576
step: 40, loss: 0.00011569978960324079
step: 50, loss: 0.00105076446197927
step: 60, loss: 7.064142846502364e-05
step: 70, loss: 0.0019493424333631992
step: 80, loss: 0.017612405121326447
step: 90, loss: 0.0008727236418053508
step: 100, loss: 0.00031799153657630086
step: 110, loss: 0.0845201313495636
step: 120, loss: 0.0002563423477113247
step: 130, loss: 0.03781285136938095
step: 140, loss: 0.0037033045664429665
step: 150, loss: 2.3014526959741488e-05
step: 160, loss: 0.0029783432837575674
step: 170, loss: 0.008894745260477066
step: 180, loss: 0.027778249233961105
step: 190, loss: 0.004873455036431551
step: 200, loss: 0.004014516714960337
step: 210, loss: 0.0012892077211290598
/usr/local/lib/python3.7/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=0.5992217898832685, f1=0.6372745490981965,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020946508273482323
step: 10, loss: 0.011694247834384441
step: 20, loss: 0.00035606196615844965
```

```
step: 30, loss: 0.0019626696594059467
step: 40, loss: 0.007676017936319113
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 50, loss: 0.08571197837591171
step: 60, loss: 0.005697306245565414
step: 70, loss: 0.0011869517620652914
step: 80, loss: 0.010251614265143871
step: 90, loss: 0.006225980818271637
step: 100, loss: 0.028480619192123413
step: 110, loss: 0.012069669552147388
step: 120, loss: 0.0016839948948472738
step: 130, loss: 0.0011877674842253327
step: 140, loss: 0.0015857057878747582
step: 150, loss: 0.0015402461867779493
step: 160, loss: 0.012476367875933647
step: 170, loss: 0.006307450123131275
step: 180, loss: 0.14485760033130646
step: 190, loss: 0.013125729747116566
step: 200, loss: 0.02843763679265976
step: 210, loss: 0.0029360793996602297
/usr/local/lib/python3.7/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=0.5898989898989899, f1=0.6428571428571429,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001343014882877469
step: 10, loss: 0.0319172702729702
step: 20, loss: 0.0023971544578671455
step: 30, loss: 0.0010640745749697709
step: 40, loss: 0.0036943326704204082
step: 50, loss: 0.02625591866672039
step: 60, loss: 0.04484853520989418
step: 70, loss: 0.0025480345357209444
step: 80, loss: 0.015491017140448093
step: 90, loss: 0.0001875185698736459
step: 100, loss: 0.0003749909810721874
step: 110, loss: 0.0011565984459593892
step: 120, loss: 0.0004703746235463768
step: 130, loss: 0.015179387293756008
step: 140, loss: 0.00451760645955801
step: 150, loss: 0.025048455223441124
step: 160, loss: 0.0006053715478628874
step: 170, loss: 0.028409423306584358
step: 180, loss: 0.0181985255330801
step: 190, loss: 0.02762831188738346
```

```
step: 200, loss: 0.0001932932500494644
step: 210, loss: 0.0018012297805398703
/usr/local/lib/python3.7/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=0.5971370143149284, f1=0.6386554621848739,
best_f1=0.5877192982456141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013982984237372875
step: 10, loss: 0.0006960671744309366
step: 20, loss: 0.014431110583245754
step: 30, loss: 0.08204904198646545
step: 40, loss: 0.0013770624063909054
step: 50, loss: 0.00011474275379441679
step: 60, loss: 0.0011499824468046427
step: 70, loss: 0.0045915930531919
step: 80, loss: 0.0007262536091729999
step: 90, loss: 0.0007667383179068565
step: 100, loss: 0.0028952087741345167
step: 110, loss: 0.045291174203157425
step: 120, loss: 0.003608901286497712
step: 130, loss: 0.000711047207005322
step: 140, loss: 0.002296398626640439
step: 150, loss: 0.003214847994968295
step: 160, loss: 0.004258488770574331
step: 170, loss: 0.007221555802971125
step: 180, loss: 0.0007385880453512073
step: 190, loss: 0.0017506186850368977
step: 200, loss: 0.0006143439095467329
step: 210, loss: 0.04924590885639191
/usr/local/lib/python3.7/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=0.5836909871244635, f1=0.6136865342163355,
best_f1=0.5877192982456141
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, 512.70it/s]
load_f1 = 0.608695652173913
real_f1 = 0.6024096385542169
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 205.65it/s]
```

3.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

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

details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4832531213760376
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5119920969009399
step: 20, loss: 0.2767992317676544
step: 30, loss: 0.3453461229801178
step: 40, loss: 0.2362372726202011
step: 50, loss: 0.30759289860725403
step: 60, loss: 0.4494229555130005
step: 70, loss: 0.47234809398651123
step: 80, loss: 0.17763851583003998
step: 90, loss: 0.3088090121746063
step: 100, loss: 0.42142701148986816
step: 110, loss: 0.24366483092308044
step: 120, loss: 0.39590027928352356
step: 130, loss: 0.30721285939216614
step: 140, loss: 0.15565913915634155
step: 150, loss: 0.30075204372406006
step: 160, loss: 0.23360830545425415
step: 170, loss: 0.38804835081100464
step: 180, loss: 0.15743067860603333
step: 190, loss: 0.16142530739307404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3831007182598114
step: 10, loss: 0.35264161229133606
step: 20, loss: 0.5436238050460815
step: 30, loss: 0.27252161502838135
step: 40, loss: 0.5958604216575623
step: 50, loss: 0.29625385999679565
step: 60, loss: 0.4875333309173584
step: 70, loss: 0.32174360752105713
step: 80, loss: 0.15141522884368896
step: 90, loss: 0.3129728436470032
step: 100, loss: 0.2554641664028168

```
step: 110, loss: 0.3676399886608124
step: 120, loss: 0.2514161169528961
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.560350239276886
step: 140, loss: 0.30933672189712524
step: 150, loss: 0.32602059841156006
step: 160, loss: 0.31393542885780334
step: 170, loss: 0.24243983626365662
step: 180, loss: 0.17359858751296997
step: 190, loss: 0.23208899796009064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37791332602500916
step: 10, loss: 0.38173729181289673
step: 20, loss: 0.45254889130592346
step: 30, loss: 0.30994662642478943
step: 40, loss: 0.08019829541444778
step: 50, loss: 0.3786347806453705
step: 60, loss: 0.16230237483978271
step: 70, loss: 0.3841885030269623
step: 80, loss: 0.31200459599494934
step: 90, loss: 0.37156569957733154
step: 100, loss: 0.5257407426834106
step: 110, loss: 0.6719605326652527
step: 120, loss: 0.368972122266921997
step: 130, loss: 0.16509072482585907
step: 140, loss: 0.3940308392047882
step: 150, loss: 0.33168166875839233
step: 160, loss: 0.6311963796615601
step: 170, loss: 0.42672592401504517
step: 180, loss: 0.3766343891620636
step: 190, loss: 0.15035001933574677
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23096832633018494
step: 10, loss: 0.2463098019361496
```

```
step: 20, loss: 0.31044402718544006
step: 30, loss: 0.2441243976354599
step: 40, loss: 0.5537114143371582
step: 50, loss: 0.24613916873931885
step: 60, loss: 0.3751182556152344
step: 70, loss: 0.30014991760253906
step: 80, loss: 0.17552898824214935
step: 90, loss: 0.2463558316230774
step: 100, loss: 0.3243370056152344
step: 110, loss: 0.3645634055137634
step: 120, loss: 0.1344311535358429
step: 130, loss: 0.3528299629688263
step: 140, loss: 0.3012997806072235
step: 150, loss: 0.12233508378267288
step: 160, loss: 0.17781642079353333
step: 170, loss: 0.22475476562976837
step: 180, loss: 0.17564354836940765
step: 190, loss: 0.0868191346526146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7020202020202021, f1=0.7004830917874396,
best_f1=0.7004830917874396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07154519855976105
step: 10, loss: 0.29803115129470825
step: 20, loss: 0.08188881725072861
step: 30, loss: 0.07974110543727875
step: 40, loss: 0.15986348688602448
step: 50, loss: 0.11888027936220169
step: 60, loss: 0.04347608983516693
step: 70, loss: 0.07076222449541092
step: 80, loss: 0.1820962131023407
step: 90, loss: 0.17869041860103607
step: 100, loss: 0.18413563072681427
step: 110, loss: 0.217031791806221
step: 120, loss: 0.0768369510769844
step: 130, loss: 0.47310134768486023
step: 140, loss: 0.12818631529808044
step: 150, loss: 0.19540616869926453
step: 160, loss: 0.09373638778924942
step: 170, loss: 0.06187507137656212
```

```
step: 180, loss: 0.04039432108402252
step: 190, loss: 0.14072000980377197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8049382716049381, f1=0.7777777777777778,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14848075807094574
step: 10, loss: 0.20034913718700041
step: 20, loss: 0.0484318770468235
step: 30, loss: 0.32235050201416016
step: 40, loss: 0.11406104266643524
step: 50, loss: 0.1470026820898056
step: 60, loss: 0.06864172220230103
step: 70, loss: 0.23641376197338104
step: 80, loss: 0.2084658443927765
step: 90, loss: 0.05028747394680977
step: 100, loss: 0.19302086532115936
step: 110, loss: 0.0160352885723114
step: 120, loss: 0.05160253494977951
step: 130, loss: 0.3191012442111969
step: 140, loss: 0.07752183824777603
step: 150, loss: 0.041870612651109695
step: 160, loss: 0.18386149406433105
step: 170, loss: 0.27718809247016907
step: 180, loss: 0.027833791449666023
step: 190, loss: 0.0750332623720169
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8279569892473119, f1=0.8186813186813187,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08225253224372864
step: 10, loss: 0.04289814829826355
```

```
step: 20, loss: 0.00817089807242155
step: 30, loss: 0.12150492519140244
step: 40, loss: 0.06699442863464355
step: 50, loss: 0.033042483031749725
step: 60, loss: 0.018057750537991524
step: 70, loss: 0.01070515438914299
step: 80, loss: 0.009753973223268986
step: 90, loss: 0.029489856213331223
step: 100, loss: 0.2819838225841522
step: 110, loss: 0.3414948582649231
step: 120, loss: 0.08479040861129761
step: 130, loss: 0.03838353976607323
step: 140, loss: 0.11259651929140091
step: 150, loss: 0.013558120466768742
step: 160, loss: 0.16119378805160522
step: 170, loss: 0.09671792387962341
step: 180, loss: 0.021636933088302612
step: 190, loss: 0.13695605099201202
/usr/local/lib/python3.7/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=0.8042328042328043, f1=0.8082901554404145,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11262288689613342
step: 10, loss: 0.06388566642999649
step: 20, loss: 0.006702039856463671
step: 30, loss: 0.17482693493366241
step: 40, loss: 0.0682055652141571
step: 50, loss: 0.03070494346320629
step: 60, loss: 0.1242787167429924
step: 70, loss: 0.0035302797332406044
step: 80, loss: 0.47338131070137024
step: 90, loss: 0.02216443046927452
step: 100, loss: 0.056525830179452896
step: 110, loss: 0.0161492470651865
step: 120, loss: 0.002682093530893326
step: 130, loss: 0.021695926785469055
step: 140, loss: 0.073259636759758
step: 150, loss: 0.04513338953256607
step: 160, loss: 0.11132469028234482
step: 170, loss: 0.002952558221295476
step: 180, loss: 0.2702271342277527
step: 190, loss: 0.030189527198672295
/usr/local/lib/python3.7/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=0.8074866310160429, f1=0.8065395095367848,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0278047211468219976
step: 10, loss: 0.0013862416381016374
step: 20, loss: 0.018086407333612442
step: 30, loss: 0.013212667778134346
step: 40, loss: 0.13575667142868042
step: 50, loss: 0.035478558391332626
step: 60, loss: 0.012928644195199013
step: 70, loss: 0.007576887030154467
step: 80, loss: 0.15070396661758423
step: 90, loss: 0.03629571944475174
step: 100, loss: 0.021997883915901184
step: 110, loss: 0.010516700334846973
step: 120, loss: 0.13349124789237976
step: 130, loss: 0.06827542185783386
step: 140, loss: 0.03720570728182793
step: 150, loss: 0.0027991486713290215
step: 160, loss: 0.004548870958387852
step: 170, loss: 0.09351958334445953
step: 180, loss: 0.058383744210004807
step: 190, loss: 0.004090309143066406
/usr/local/lib/python3.7/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=0.8140161725067385, f1=0.8306010928961748,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017817243933677673
step: 10, loss: 0.027917729690670967
step: 20, loss: 0.008557136170566082
step: 30, loss: 0.029744215309619904
step: 40, loss: 0.0015251114964485168
step: 50, loss: 0.007385988254100084
step: 60, loss: 0.0007658089743927121
step: 70, loss: 0.0007849642424844205
step: 80, loss: 0.04040229693055153
step: 90, loss: 0.03710344806313515
step: 100, loss: 0.019424255937337875
step: 110, loss: 0.001303168130107224
step: 120, loss: 0.002187657868489623
step: 130, loss: 0.00563295790925622
```

```
step: 140, loss: 0.07855138182640076
step: 150, loss: 0.0028278776444494724
step: 160, loss: 0.005795991513878107
step: 170, loss: 0.0020089654717594385
step: 180, loss: 0.028212212026119232
step: 190, loss: 0.0016109615098685026
/usr/local/lib/python3.7/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=0.8043478260869565, f1=0.8346883468834688,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01798371784389019
step: 10, loss: 0.0006830818601883948
step: 20, loss: 0.013740822672843933
step: 30, loss: 0.009558087214827538
step: 40, loss: 0.00048784638056531555
step: 50, loss: 0.012993474490094057
step: 60, loss: 0.0009616040624678135
step: 70, loss: 0.005440490785986185
step: 80, loss: 0.011141580529510975
step: 90, loss: 0.0025901058688759804
step: 100, loss: 0.0008582512382417917
step: 110, loss: 0.05310821160674095
step: 120, loss: 0.0007663594442419708
step: 130, loss: 0.00332559272646904
step: 140, loss: 0.0011986554600298405
step: 150, loss: 0.0008385584224015474
step: 160, loss: 0.001116268802434206
step: 170, loss: 0.03244611993432045
step: 180, loss: 0.021551713347434998
step: 190, loss: 0.0041731311939656734
/usr/local/lib/python3.7/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=0.7968337730870713, f1=0.8432432432432432,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00117722328286618
step: 10, loss: 0.010859301313757896
step: 20, loss: 0.005481306463479996
step: 30, loss: 0.03954877331852913
step: 40, loss: 0.0025859600864350796
step: 50, loss: 0.00801783986389637
```

```
step: 60, loss: 0.05811955779790878
step: 70, loss: 0.0033983560279011726
step: 80, loss: 0.00453920429572463
step: 90, loss: 0.010692216455936432
step: 100, loss: 0.0008776510367169976
step: 110, loss: 0.3186434507369995
step: 120, loss: 0.0016384981572628021
step: 130, loss: 0.0011628203792497516
step: 140, loss: 0.014309865422546864
step: 150, loss: 0.016047868877649307
step: 160, loss: 0.0006484594196081161
step: 170, loss: 0.0039710537530481815
step: 180, loss: 0.0005595095572061837
step: 190, loss: 0.006430821027606726
/usr/local/lib/python3.7/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=0.8074866310160429, f1=0.8369565217391304,
best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24165138602256775
step: 10, loss: 0.003464827546849847
step: 20, loss: 0.008450162597000599
step: 30, loss: 0.13859426975250244
step: 40, loss: 0.0005524446605704725
step: 50, loss: 0.0007176481885835528
step: 60, loss: 0.005481051746755838
step: 70, loss: 0.05073659494519234
step: 80, loss: 0.00124901975505054
step: 90, loss: 0.005512874573469162
step: 100, loss: 0.001176379737444222
step: 110, loss: 0.0015380746917799115
step: 120, loss: 0.025377320125699043
step: 130, loss: 0.0010233588982373476
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.013958985917270184
step: 150, loss: 0.0009308847947977483
step: 160, loss: 0.005936883855611086
step: 170, loss: 0.0002745770907495171
step: 180, loss: 0.0012910474324598908
step: 190, loss: 0.0609777607023716
/usr/local/lib/python3.7/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=0.7955182072829131, f1=0.8347338935574229,
best_f1=0.8186813186813187
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06473159044981003
step: 10, loss: 0.013987851329147816
step: 20, loss: 0.003996615298092365
step: 30, loss: 0.017828693613409996
step: 40, loss: 0.00029106930014677346
step: 50, loss: 0.010670140385627747
step: 60, loss: 0.007007171865552664
step: 70, loss: 0.00023502780823037028
step: 80, loss: 0.00041752413380891085
step: 90, loss: 0.001298540155403316
step: 100, loss: 0.017233174294233322
step: 110, loss: 0.007064176723361015
step: 120, loss: 0.0004312056698836386
step: 130, loss: 0.0007424113573506474
step: 140, loss: 0.003056894987821579
step: 150, loss: 0.007355473469942808
step: 160, loss: 0.042276266696228981
step: 170, loss: 0.00028545004897750914
step: 180, loss: 0.0004938391502946615
step: 190, loss: 0.004375978838652372
/usr/local/lib/python3.7/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=0.796875, f1=0.8222811671087534, best_f1=0.8186813186813187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006367877591401339
step: 10, loss: 0.0003566014929674566
step: 20, loss: 0.009296223521232605
step: 30, loss: 0.0005271699628792703
step: 40, loss: 0.0012062422465533018
step: 50, loss: 0.0003328302700538188
step: 60, loss: 0.00045662609045393765
step: 70, loss: 0.0016062961658462882
step: 80, loss: 0.0007938978378660977
step: 90, loss: 0.003801333485171199
step: 100, loss: 0.00032737827859818935
step: 110, loss: 0.00651979073882103
step: 120, loss: 0.0009625912643969059
step: 130, loss: 0.00034171828883700073
step: 140, loss: 0.000958646007347852
step: 150, loss: 0.0002675323630683124
step: 160, loss: 0.1924063265323639
step: 170, loss: 0.11859031766653061
```

step: 180, loss: 0.0013963356614112854
step: 190, loss: 0.0032779155299067497
/usr/local/lib/python3.7/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=0.8011204481792716, f1=0.8459383753501402,
best_f1=0.8186813186813187
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, 249.75it/s]
load_f1 = 0.8277777777777778
real_f1 = 0.8360655737704918
733it [00:00, 2525.09it/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, 196.44it/s]

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

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


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

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

Selected optimization level 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

313

```
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: 0.49499186873435974
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.626232385635376
step: 20, loss: 0.2892874777317047
step: 30, loss: 0.48398247361183167
step: 40, loss: 0.5863659381866455
step: 50, loss: 0.33483678102493286
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.5672401189804077
step: 70, loss: 0.3156864047050476
step: 80, loss: 0.2766721248626709
step: 90, loss: 0.17993861436843872
step: 100, loss: 0.2366352528333664
step: 110, loss: 0.4217384159564972
step: 120, loss: 0.3228040039539337
step: 130, loss: 0.31681394577026367
step: 140, loss: 0.38358160853385925
step: 150, loss: 0.3171163499355316
step: 160, loss: 0.39186209440231323
step: 170, loss: 0.29756584763526917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45730027548209373, f1=0.3812316715542522,
best_f1=0.3812316715542522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27907702326774597
step: 10, loss: 0.42372003197669983
step: 20, loss: 0.24966196715831757
step: 30, loss: 0.19123037159442902
step: 40, loss: 0.07691429555416107
step: 50, loss: 0.30454930663108826
step: 60, loss: 0.18143607676029205
step: 70, loss: 0.15796318650245667
step: 80, loss: 0.15080314874649048
step: 90, loss: 0.10490627586841583
step: 100, loss: 0.19264569878578186
step: 110, loss: 0.13401825726032257
step: 120, loss: 0.13266277313232422
step: 130, loss: 0.2935445010662079
step: 140, loss: 0.30929073691368103
step: 150, loss: 0.15227238833904266
step: 160, loss: 0.16343598067760468
step: 170, loss: 0.038486726582050323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7411444141689374, f1=0.7373737373737375,
best_f1=0.7373737373737375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2387998253107071
step: 10, loss: 0.026272110641002655
step: 20, loss: 0.07232926040887833
step: 30, loss: 0.017413340508937836
step: 40, loss: 0.16379980742931366
step: 50, loss: 0.24788856506347656
step: 60, loss: 0.012488861568272114
step: 70, loss: 0.10808929800987244
step: 80, loss: 0.17000123858451843
```

```
step: 90, loss: 0.3666585683822632
step: 100, loss: 0.13825459778308868
step: 110, loss: 0.006149623077362776
step: 120, loss: 0.0840306356549263
step: 130, loss: 0.18980641663074493
step: 140, loss: 0.15700387954711914
step: 150, loss: 0.040229398757219315
step: 160, loss: 0.06355153769254684
step: 170, loss: 0.02577603980898857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.78125, f1=0.8058252427184465, best_f1=0.8058252427184465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06018939986824989
step: 10, loss: 0.0760040283203125
step: 20, loss: 0.005804680287837982
step: 30, loss: 0.15216149389743805
step: 40, loss: 0.03842591121792793
step: 50, loss: 0.07005060464143753
step: 60, loss: 0.10612853616476059
step: 70, loss: 0.009908273816108704
step: 80, loss: 0.24243436753749847
step: 90, loss: 0.27540528774261475
step: 100, loss: 0.19389308989048004
step: 110, loss: 0.285308301448822
step: 120, loss: 0.3904467225074768
step: 130, loss: 0.042779237031936646
step: 140, loss: 0.06730514019727707
step: 150, loss: 0.23470419645309448
step: 160, loss: 0.007729369215667248
step: 170, loss: 0.24044226109981537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7959183673469389, f1=0.8382352941176471,
best_f1=0.8382352941176471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029908858239650726
step: 10, loss: 0.10703904181718826
step: 20, loss: 0.02243657037615776
step: 30, loss: 0.010406131856143475
step: 40, loss: 0.038260720670223236
step: 50, loss: 0.01675361581146717
step: 60, loss: 0.024785371497273445
step: 70, loss: 0.04294576868414879
step: 80, loss: 0.004563279915601015
step: 90, loss: 0.06360418349504471
step: 100, loss: 0.011817374266684055
step: 110, loss: 0.06906218081712723
step: 120, loss: 0.12061109393835068
step: 130, loss: 0.0027316310442984104
step: 140, loss: 0.027475010603666306
step: 150, loss: 0.07629161328077316
step: 160, loss: 0.17767558991909027
step: 170, loss: 0.023413890972733498
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8075880758807589, f1=0.8346456692913385,
best_f1=0.8346456692913385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09358706325292587
step: 10, loss: 0.023619884625077248
step: 20, loss: 0.09783083200454712
step: 30, loss: 0.0027119116857647896
step: 40, loss: 0.004070433788001537
step: 50, loss: 0.010361041873693466
step: 60, loss: 0.02216942422091961
step: 70, loss: 0.032478656619787216
step: 80, loss: 0.009965198114514351
step: 90, loss: 0.1839073896408081
step: 100, loss: 0.01332907099276781
step: 110, loss: 0.10713419318199158
step: 120, loss: 0.017692573368549347
step: 130, loss: 0.13409897685050964
step: 140, loss: 0.08768502622842789
step: 150, loss: 0.03494533151388168
```

```
step: 160, loss: 0.07364556193351746
step: 170, loss: 0.00582149438560009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8115183246073299, f1=0.8492462311557789,
best_f1=0.8492462311557789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1273476928472519
step: 10, loss: 0.03591534495353699
step: 20, loss: 0.012351498939096928
step: 30, loss: 0.27787908911705017
step: 40, loss: 0.0037847491912543774
step: 50, loss: 0.006042980123311281
step: 60, loss: 0.05230169743299484
step: 70, loss: 0.003711032448336482
step: 80, loss: 0.05521383881568909
step: 90, loss: 0.047715939581394196
step: 100, loss: 0.004916031379252672
step: 110, loss: 0.02318611741065979
step: 120, loss: 0.02668171375989914
step: 130, loss: 0.003762622596696019
step: 140, loss: 0.01904258131980896
step: 150, loss: 0.14979392290115356
step: 160, loss: 0.014540309086441994
step: 170, loss: 0.03097653016448021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8118279569892471, f1=0.8708860759493672,
best_f1=0.8708860759493672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031700603663921356
step: 10, loss: 0.005185144022107124
step: 20, loss: 0.017164641991257668
step: 30, loss: 0.000473497697385028
```

```
step: 40, loss: 0.016166573390364647
step: 50, loss: 0.008064914494752884
step: 60, loss: 0.029212534427642822
step: 70, loss: 0.048531875014305115
step: 80, loss: 0.0035089042503386736
step: 90, loss: 0.033082135021686554
step: 100, loss: 0.0018085968913510442
step: 110, loss: 0.062003012746572495
step: 120, loss: 0.04046221077442169
step: 130, loss: 0.00945042073726654
step: 140, loss: 0.03414839506149292
step: 150, loss: 0.0008414333569817245
step: 160, loss: 0.0019824428018182516
step: 170, loss: 0.09268281608819962
/usr/local/lib/python3.7/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=0.8066298342541437, f1=0.8519480519480519,
best_f1=0.8708860759493672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015673788264393806
step: 10, loss: 0.023334329947829247
step: 20, loss: 0.005362370982766151
step: 30, loss: 0.1895218789577484
step: 40, loss: 0.24162644147872925
step: 50, loss: 0.013280740939080715
step: 60, loss: 0.038172196596860886
step: 70, loss: 0.13220761716365814
step: 80, loss: 0.017517831176519394
step: 90, loss: 0.03724779188632965
step: 100, loss: 0.010295557789504528
step: 110, loss: 0.07075303047895432
step: 120, loss: 0.017471229657530785
step: 130, loss: 0.02483006753027439
step: 140, loss: 0.20961663126945496
step: 150, loss: 0.25348523259162903
step: 160, loss: 0.0043502855114638805
step: 170, loss: 0.03265528380870819
/usr/local/lib/python3.7/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=0.8108108108108109, f1=0.8505154639175257,
best_f1=0.8708860759493672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.011407606303691864
step: 10, loss: 0.004603066947311163
step: 20, loss: 0.17566649615764618
step: 30, loss: 0.02295491471886635
step: 40, loss: 0.009752042591571808
step: 50, loss: 0.14523613452911377
step: 60, loss: 0.033617809414863586
step: 70, loss: 0.009703894145786762
step: 80, loss: 0.029298510402441025
step: 90, loss: 0.003853282891213894
step: 100, loss: 0.013632369227707386
step: 110, loss: 0.0038778777234256268
step: 120, loss: 0.018958136439323425
step: 130, loss: 0.0455927699804306
step: 140, loss: 0.03142537549138069
step: 150, loss: 0.0022268800530582666
step: 160, loss: 0.0013895524898543954
step: 170, loss: 0.02222578041255474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8160000000000001, f1=0.8578811369509044,
best_f1=0.8578811369509044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006164085119962692
step: 10, loss: 0.0008221528260037303
step: 20, loss: 0.0003987761738244444
step: 30, loss: 0.006225171498954296
step: 40, loss: 0.008459823206067085
step: 50, loss: 0.0029001843649894
step: 60, loss: 0.025606045499444008
step: 70, loss: 0.0009288354776799679
step: 80, loss: 0.05367443338036537
step: 90, loss: 0.0189653467386961
step: 100, loss: 0.06366772949695587
step: 110, loss: 0.0009933365508913994
step: 120, loss: 0.0023681065067648888
step: 130, loss: 0.017515327781438828
step: 140, loss: 0.006732830312103033
step: 150, loss: 0.0005045410944148898
step: 160, loss: 0.01876200921833515
step: 170, loss: 0.08690519630908966
```

```
/usr/local/lib/python3.7/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=0.8128078817733989, f1=0.8345323741007193,
best_f1=0.8578811369509044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011217205319553614
step: 10, loss: 0.0050602261908352375
step: 20, loss: 0.04368007928133011
step: 30, loss: 0.038021404296159744
step: 40, loss: 0.013962388038635254
step: 50, loss: 0.02130727283656597
step: 60, loss: 0.003973803017288446
step: 70, loss: 0.00022789309150539339
step: 80, loss: 0.00029380395426414907
step: 90, loss: 0.00731619680300355
step: 100, loss: 0.0012325822608545423
step: 110, loss: 0.032701313495635986
step: 120, loss: 0.015368612483143806
step: 130, loss: 0.0507420189678669
step: 140, loss: 0.0008185969782061875
step: 150, loss: 0.0027057169936597347
step: 160, loss: 0.021276485174894333
step: 170, loss: 0.03659756854176521
/usr/local/lib/python3.7/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=0.8064516129032259, f1=0.8622448979591837,
best_f1=0.8578811369509044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002752444474026561
step: 10, loss: 0.00023633410455659032
step: 20, loss: 0.00029110285686329007
step: 30, loss: 0.03105062060058117
step: 40, loss: 0.0009741219691932201
step: 50, loss: 0.0639089047908783
step: 60, loss: 0.02451452612876892
step: 70, loss: 0.12423151731491089
step: 80, loss: 0.000824520131573081
step: 90, loss: 0.005041879136115313
step: 100, loss: 0.018166866153478622
step: 110, loss: 0.004861972760409117
step: 120, loss: 0.009688653983175755
step: 130, loss: 0.0005955502274446189
```

```
step: 140, loss: 0.029294095933437347
step: 150, loss: 0.00046750763431191444
step: 160, loss: 0.00029743817867711186
step: 170, loss: 0.000261227396549657
/usr/local/lib/python3.7/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=0.8044077134986227, f1=0.8465608465608466,
best_f1=0.8578811369509044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0052403854206204414
step: 10, loss: 0.00022258779790718108
step: 20, loss: 0.018917428329586983
step: 30, loss: 0.0005842004320584238
step: 40, loss: 0.00041961262468248606
step: 50, loss: 0.00272948807105422
step: 60, loss: 0.010322825983166695
step: 70, loss: 0.02598845399916172
step: 80, loss: 0.006977694109082222
step: 90, loss: 0.0002492184576112777
step: 100, loss: 0.0009346186416223645
step: 110, loss: 0.028680190443992615
step: 120, loss: 0.0007392641273327172
step: 130, loss: 0.0009472044766880572
step: 140, loss: 0.04047619551420212
step: 150, loss: 0.013345872983336449
step: 160, loss: 0.000988535350188613
step: 170, loss: 0.0451468862593174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8181818181818182, f1=0.8556962025316455,
best_f1=0.8556962025316455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001356764551019296
step: 10, loss: 0.23358295857906342
step: 20, loss: 0.001238205935806036
step: 30, loss: 0.01720374822616577
step: 40, loss: 0.0016710925847291946
step: 50, loss: 0.0016356317792087793
```

```
step: 60, loss: 0.002846557879820466
step: 70, loss: 0.04875403642654419
step: 80, loss: 0.0003732536861207336
step: 90, loss: 0.005738475359976292
step: 100, loss: 0.027048790827393532
step: 110, loss: 0.0025601566303521395
step: 120, loss: 0.00036830120370723307
step: 130, loss: 0.001626285258680582
step: 140, loss: 0.004095795564353466
step: 150, loss: 0.0007925917161628604
step: 160, loss: 0.00035431113792583346
step: 170, loss: 0.0010661401320248842
/usr/local/lib/python3.7/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=0.8148148148148149, f1=0.8492462311557789,
best_f1=0.8556962025316455
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:06, 318.47it/s]
```

```
load_f1 = 0.5344827586206897
real_f1 = 0.5073170731707317
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, 196.41it/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 roberta \
  --fp16 \
  --save_model \
  --dk general  \
  --dk product \
  --summarize


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

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

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

```
Defaults for this optimization level are:
enabled                  : True
opt_level                : O2
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                  : True
opt_level                : O2
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.639726996421814
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44198018312454224
step: 20, loss: 0.45858070254325867
step: 30, loss: 0.28914451599121094
step: 40, loss: 0.36573952436447144
step: 50, loss: 0.6341001391410828
step: 60, loss: 0.40501052141189575
step: 70, loss: 0.15148121118545532
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.17763565480709076
step: 90, loss: 0.18001675605773926
step: 100, loss: 0.1334400773048401
step: 110, loss: 0.08209965378046036
step: 120, loss: 0.036432910710573196
step: 130, loss: 0.025756636634469032
step: 140, loss: 0.05366101488471031
step: 150, loss: 0.21993540227413177
step: 160, loss: 0.1078152060508728
step: 170, loss: 0.14024218916893005
step: 180, loss: 0.0810016542673111
step: 190, loss: 0.048809681087732315
step: 200, loss: 0.019278554245829582
step: 210, loss: 0.03701651468873024
```

```
step: 220, loss: 0.10395180433988571
step: 230, loss: 0.002539142034947872
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9524861878453039, f1=0.962962962962963,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003664445597678423
step: 10, loss: 0.20918084681034088
step: 20, loss: 0.015576707199215889
step: 30, loss: 0.022966332733631134
step: 40, loss: 0.01732923649251461
step: 50, loss: 0.023230943828821182
step: 60, loss: 0.025028858333826065
step: 70, loss: 0.02558201737701893
step: 80, loss: 0.003333626314997673
step: 90, loss: 0.021699467673897743
step: 100, loss: 0.0028282764833420515
step: 110, loss: 0.07663662731647491
step: 120, loss: 0.008837919682264328
step: 130, loss: 0.03261891007423401
step: 140, loss: 0.012712604366242886
step: 150, loss: 0.09826596826314926
step: 160, loss: 0.07950758188962936
step: 170, loss: 0.003476355690509081
step: 180, loss: 0.010040759108960629
step: 190, loss: 0.024502743035554886
step: 200, loss: 0.006675581447780132
step: 210, loss: 0.015159521251916885
step: 220, loss: 0.005391485057771206
step: 230, loss: 0.021883515641093254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9605411499436302, f1=0.9557321225879684,
best_f1=0.9557321225879684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010800721123814583
step: 10, loss: 0.00654271524399519
step: 20, loss: 0.13599224388599396
step: 30, loss: 0.023112913593649864
step: 40, loss: 0.038447000086307526
step: 50, loss: 0.010172228328883648
step: 60, loss: 0.007861468009650707
step: 70, loss: 0.006181090604513884
step: 80, loss: 0.028525609523057938
step: 90, loss: 0.00931596476584673
step: 100, loss: 0.0015387069433927536
step: 110, loss: 0.01880962960422039
step: 120, loss: 0.007781077176332474
step: 130, loss: 0.08268902450799942
step: 140, loss: 0.0015830837655812502
step: 150, loss: 0.10930224508047104
step: 160, loss: 0.0013312004739418626
step: 170, loss: 0.01751548983156681
step: 180, loss: 0.07144492119550705
step: 190, loss: 0.009222621098160744
step: 200, loss: 0.08449448645114899
step: 210, loss: 0.009256311692297459
step: 220, loss: 0.09638246893882751
step: 230, loss: 0.011427175253629684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9622222222222223, f1=0.9648924122310306,
best_f1=0.9648924122310306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02770005166530609
step: 10, loss: 0.005957509856671095
step: 20, loss: 0.03643631562590599
step: 30, loss: 0.0012681498192250729
step: 40, loss: 0.05032789707183838
step: 50, loss: 0.0263417586684227
step: 60, loss: 0.002632661024108529
step: 70, loss: 0.24700401723384857
step: 80, loss: 0.01826883666217327
step: 90, loss: 0.04599764198064804
```

```
step: 100, loss: 0.09131790697574615
step: 110, loss: 0.0675995722413063
step: 120, loss: 0.003540464909747243
step: 130, loss: 0.0073833223432302475
step: 140, loss: 0.011225392110645771
step: 150, loss: 0.002525013405829668
step: 160, loss: 0.005878155119717121
step: 170, loss: 0.05034773424267769
step: 180, loss: 0.18035924434661865
step: 190, loss: 0.008079416118562222
step: 200, loss: 0.13784591853618622
step: 210, loss: 0.004819610621780157
step: 220, loss: 0.00163117831107229
step: 230, loss: 0.005861918441951275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.972129319955407, f1=0.9665178571428571,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03264149650931358
step: 10, loss: 0.019735289737582207
step: 20, loss: 0.03939807787537575
step: 30, loss: 0.00154300006897076964
step: 40, loss: 0.0013397971633821726
step: 50, loss: 0.002765567507594824
step: 60, loss: 0.03333912044763565
step: 70, loss: 0.003014859277755022
step: 80, loss: 0.0303370151668787
step: 90, loss: 0.020949792116880417
step: 100, loss: 0.02185797318816185
step: 110, loss: 0.0033636733423918486
step: 120, loss: 0.0014372188597917557
step: 130, loss: 0.0004655668162740767
step: 140, loss: 0.003089895471930504
step: 150, loss: 0.04029844328761101
step: 160, loss: 0.006823661737143993
step: 170, loss: 0.08970937132835388
step: 180, loss: 0.004158985801041126
step: 190, loss: 0.016986748203635216
step: 200, loss: 0.011677847243845463
step: 210, loss: 0.0019971742294728756
```

```
step: 220, loss: 0.00962473638355732
step: 230, loss: 0.008908829651772976
/usr/local/lib/python3.7/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=0.9694915254237287, f1=0.971687429218573,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009468800271861255
step: 10, loss: 0.002319120801985264
step: 20, loss: 0.011546689085662365
step: 30, loss: 0.0006694787298329175
step: 40, loss: 0.00022466902737505734
step: 50, loss: 0.00288152857683599
step: 60, loss: 0.0008294770959764719
step: 70, loss: 0.12583455443382263
step: 80, loss: 0.0005729185650125146
step: 90, loss: 0.009258112870156765
step: 100, loss: 0.0027653861325234175
step: 110, loss: 0.00771584315225482
step: 120, loss: 0.0021411608904600143
step: 130, loss: 0.0013058183249086142
step: 140, loss: 0.0014970776392146945
step: 150, loss: 0.0005705817602574825
step: 160, loss: 0.0007497944170609117
step: 170, loss: 0.0003068714286200702
step: 180, loss: 0.00029265740886330605
step: 190, loss: 0.0006765344878658652
step: 200, loss: 0.005874320864677429
step: 210, loss: 0.0040602535009384155
step: 220, loss: 0.04665665328502655
step: 230, loss: 0.0017737757880240679
/usr/local/lib/python3.7/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=0.9686800894854586, f1=0.9707207207207207,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005856949370354414
step: 10, loss: 0.0003611220163293183
step: 20, loss: 0.0015692978631705046
step: 30, loss: 0.012892086990177631
step: 40, loss: 0.25142034888267517
step: 50, loss: 0.04320374131202698
```

```
step: 60, loss: 0.006677234545350075
step: 70, loss: 0.004724560305476189
step: 80, loss: 0.005094637628644705
step: 90, loss: 0.0031181366648525
step: 100, loss: 0.005224430467933416
step: 110, loss: 0.00032606773311272264
step: 120, loss: 0.0016510547138750553
step: 130, loss: 0.012903185561299324
step: 140, loss: 0.0003736330254469067
step: 150, loss: 0.09725340455770493
step: 160, loss: 0.004874671809375286
step: 170, loss: 0.008157534524798393
step: 180, loss: 0.0012821451527997851
step: 190, loss: 0.0037846865598112345
step: 200, loss: 0.0060145496390759945
step: 210, loss: 0.07895655930042267
step: 220, loss: 0.004055083263665438
step: 230, loss: 0.005843461025506258
/usr/local/lib/python3.7/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=0.9664429530201343, f1=0.9599109131403119,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005178486462682486
step: 10, loss: 0.0070046489126980305
step: 20, loss: 0.0006066297646611929
step: 30, loss: 0.002720105927437544
step: 40, loss: 0.001209368696436286
step: 50, loss: 0.003987895790487528
step: 60, loss: 0.00044190671178512275
step: 70, loss: 0.0002465941070113331
step: 80, loss: 0.03218263015151024
step: 90, loss: 0.0002870571333914995
step: 100, loss: 0.01216447539627552
step: 110, loss: 0.00030732405139133334
step: 120, loss: 0.003933141008019447
step: 130, loss: 0.002186167286708951
step: 140, loss: 0.0005873502232134342
step: 150, loss: 0.23666082322597504
step: 160, loss: 0.0011041854741051793
step: 170, loss: 0.0020767436362802982
step: 180, loss: 0.001000869320705533
step: 190, loss: 0.0319240465760231
step: 200, loss: 0.008970928378403187
step: 210, loss: 0.0006187358521856368
```

```
step: 220, loss: 0.0003563020727597177
step: 230, loss: 0.00043328243191353977
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753914988814317, f1=0.9698324022346367,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029723087791353464
step: 10, loss: 0.0012084160698577762
step: 20, loss: 0.0011753879953175783
step: 30, loss: 0.001514427363872528
step: 40, loss: 0.0034625378903001547
step: 50, loss: 0.00020575470989570022
step: 60, loss: 0.00035792894777841866
step: 70, loss: 0.00023190680076368153
step: 80, loss: 4.913284647045657e-05
step: 90, loss: 0.1299659013748169
step: 100, loss: 0.00036397529765963554
step: 110, loss: 0.00027653094730339944
step: 120, loss: 0.0001973936741705984
step: 130, loss: 0.0011080713011324406
step: 140, loss: 0.0003926650097128004
step: 150, loss: 0.0003930256934836507
step: 160, loss: 0.0010499573545530438
step: 170, loss: 0.00016433846030849963
step: 180, loss: 0.0003929685626644641
step: 190, loss: 6.607421528315172e-05
step: 200, loss: 0.0045404378324747086
step: 210, loss: 0.0032435348257422447
step: 220, loss: 0.00010120005026692525
step: 230, loss: 0.000197559580556117
/usr/local/lib/python3.7/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=0.9698996655518396, f1=0.9711111111111111,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001997111365199089
step: 10, loss: 0.0002760199422482401
```

```
step: 20, loss: 0.002667937893420458
step: 30, loss: 7.338192517636344e-05
step: 40, loss: 0.000248513970291242
step: 50, loss: 5.704539216822013e-05
step: 60, loss: 7.53884669393301e-05
step: 70, loss: 0.02319882996380329
step: 80, loss: 0.0002743297372944653
step: 90, loss: 3.959783134632744e-05
step: 100, loss: 7.175716746132821e-05
step: 110, loss: 0.0038385086227208376
step: 120, loss: 0.000323760585160926
step: 130, loss: 0.0004978704382665455
step: 140, loss: 9.701668750494719e-05
step: 150, loss: 0.0005888903397135437
step: 160, loss: 5.16138861712534e-05
step: 170, loss: 9.659847273724154e-05
step: 180, loss: 0.010130243375897408
step: 190, loss: 8.84004621184431e-05
step: 200, loss: 0.0026665725745260715
step: 210, loss: 0.002086326712742448
step: 220, loss: 0.00014748242392670363
step: 230, loss: 0.00047348145744763315
/usr/local/lib/python3.7/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=0.9661399548532732, f1=0.963963963963964,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014065296272747219
step: 10, loss: 0.0009258141508325934
step: 20, loss: 4.8687430535210297e-05
step: 30, loss: 0.0003926219360437244
step: 40, loss: 3.905472840415314e-05
step: 50, loss: 0.0003092476399615407
step: 60, loss: 0.004862740635871887
step: 70, loss: 0.0004054050659760833
step: 80, loss: 0.0001383338967571035
step: 90, loss: 0.0030875965021550655
step: 100, loss: 8.975413220468909e-05
step: 110, loss: 0.00011098237155238166
step: 120, loss: 0.00011581042053876445
step: 130, loss: 9.353738278150558e-05
step: 140, loss: 0.0004087873676326126
step: 150, loss: 0.00010899986955337226
step: 160, loss: 0.06254391372203827
step: 170, loss: 0.006701661739498377
```

```
step: 180, loss: 5.7183664466720074e-05
step: 190, loss: 8.073057688307017e-05
step: 200, loss: 0.0024887537583708763
step: 210, loss: 3.7478810554603115e-05
step: 220, loss: 0.0004834742285311222
step: 230, loss: 0.0007696609827689826
/usr/local/lib/python3.7/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=0.9732739420935412, f1=0.9710467706013363,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.6173746997956187e-05
step: 10, loss: 5.0179743993794546e-05
step: 20, loss: 0.002351702656596899
step: 30, loss: 0.016039906069636345
step: 40, loss: 7.556005584774539e-05
step: 50, loss: 0.0003089364035986364
step: 60, loss: 0.131962388753891
step: 70, loss: 4.317824641475454e-05
step: 80, loss: 0.0001690334756858647
step: 90, loss: 0.003988530021160841
step: 100, loss: 3.582457793527283e-05
step: 110, loss: 0.00013792520621791482
step: 120, loss: 0.0004193284548819065
step: 130, loss: 0.00020713478443212807
step: 140, loss: 0.0004375410499051213
step: 150, loss: 5.3797673899680376e-05
step: 160, loss: 6.328710151137784e-05
step: 170, loss: 0.00014017074136063457
step: 180, loss: 7.882757927291095e-05
step: 190, loss: 0.00017382684745825827
step: 200, loss: 0.000130189408082515
step: 210, loss: 0.00010404842760181054
step: 220, loss: 3.379847476026043e-05
step: 230, loss: 9.004412277135998e-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=0.9732142857142857, f1=0.9731543624161074,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009192410507239401
step: 10, loss: 0.00017429648141842335
```

```
step: 20, loss: 5.4721254855394363e-05
step: 30, loss: 3.625280442065559e-05
step: 40, loss: 0.00014393976016435772
step: 50, loss: 0.001605628407560289
step: 60, loss: 0.00013733012019656599
step: 70, loss: 6.414009112631902e-05
step: 80, loss: 5.020013122702949e-05
step: 90, loss: 6.797272362746298e-05
step: 100, loss: 0.00011934735084651038
step: 110, loss: 4.348705988377333e-05
step: 120, loss: 6.827602919656783e-05
step: 130, loss: 5.555342067964375e-05
step: 140, loss: 0.00023182820586953312
step: 150, loss: 5.6678887631278485e-05
step: 160, loss: 0.004552455618977547
step: 170, loss: 5.34509563294705e-05
step: 180, loss: 0.03363380953669548
step: 190, loss: 4.3629002902889624e-05
step: 200, loss: 2.916069934144616e-05
step: 210, loss: 0.0028443557675927877
step: 220, loss: 4.297562554711476e-05
step: 230, loss: 5.0510734581621364e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9754464285714286, f1=0.9742441209406495,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.248303594067693e-05
step: 10, loss: 3.088538142037578e-05
step: 20, loss: 0.0005811959272250533
step: 30, loss: 4.44442848674953e-05
step: 40, loss: 6.733394548064098e-05
step: 50, loss: 4.7622077545383945e-05
step: 60, loss: 4.229476326145232e-05
step: 70, loss: 4.0000450098887086e-05
step: 80, loss: 5.777782644145191e-05
step: 90, loss: 4.8547604819759727e-05
step: 100, loss: 6.348727038130164e-05
step: 110, loss: 4.133420225116424e-05
step: 120, loss: 2.4593751732027158e-05
step: 130, loss: 0.0002597769198473543
```

```
step: 140, loss: 4.883300789515488e-05
step: 150, loss: 5.092823630548082e-05
step: 160, loss: 0.0002094627998303622
step: 170, loss: 0.00014920797548256814
step: 180, loss: 2.9868469937355258e-05
step: 190, loss: 0.0007918282644823194
step: 200, loss: 0.000218340806895867
step: 210, loss: 7.660510891582817e-05
step: 220, loss: 8.522333519067615e-05
step: 230, loss: 4.657336830860004e-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=0.9754464285714286, f1=0.9743016759776536,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003141889756079763
step: 10, loss: 4.571378667606041e-05
step: 20, loss: 0.00043786910828202963
step: 30, loss: 8.635192352812737e-05
step: 40, loss: 2.473537460900843e-05
step: 50, loss: 3.5779961763182655e-05
step: 60, loss: 3.9033104258123785e-05
step: 70, loss: 0.00012189328845124692
step: 80, loss: 0.0010329678189009428
step: 90, loss: 3.589523475966416e-05
step: 100, loss: 2.592374585219659e-05
step: 110, loss: 5.4407228162745014e-05
step: 120, loss: 0.027327684685587883
step: 130, loss: 2.933558243967127e-05
step: 140, loss: 6.797295645810664e-05
step: 150, loss: 4.695128518505953e-05
step: 160, loss: 0.000621406827121973
step: 170, loss: 3.8372101698769256e-05
step: 180, loss: 4.3088562961202115e-05
step: 190, loss: 0.00019572193559724838
step: 200, loss: 0.00017897735233418643
step: 210, loss: 0.00027856265660375357
step: 220, loss: 3.474828918115236e-05
step: 230, loss: 0.00012846826575696468
/usr/local/lib/python3.7/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=0.9754464285714286, f1=0.9743016759776536,
best_f1=0.9742441209406495
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, 212.78it/s]
load_f1 = 0.9742441209406495
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 [00:22, 195.87it/s]
```

3.3.2   DBLP-GoogleScholar - Running the matcher

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

```
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general  \
  --dk product \
  --summarize


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

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

```
Downloading: 100% 481/481 [00:00<00:00, 397kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 34.5MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 24.0MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 66.8MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```
master_weights            : True
loss_scale                : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6506847739219666
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5068556070327759
step: 20, loss: 0.3757193088531494
step: 30, loss: 0.28390127420425415
step: 40, loss: 0.29211127758026123
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.3725414276123047
step: 60, loss: 0.07296596467494965
step: 70, loss: 0.16537587344646454
step: 80, loss: 0.10842201858758926
step: 90, loss: 0.15128567814826965
step: 100, loss: 0.28047242760658264
step: 110, loss: 0.1521952897310257
step: 120, loss: 0.0984153151512146
step: 130, loss: 0.06992834806442261
step: 140, loss: 0.3529106080532074
step: 150, loss: 0.0717058777809143
step: 160, loss: 0.17735762894153595
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.574860155582428
step: 180, loss: 0.097596175968647
step: 190, loss: 0.055591437965631485
step: 200, loss: 0.0840011015534401
step: 210, loss: 0.03764836862683296
step: 220, loss: 0.08311715722084045
step: 230, loss: 0.24443088471889496
step: 240, loss: 0.038908980786800385
step: 250, loss: 0.0356547012925148
step: 260, loss: 0.2961648106575012
step: 270, loss: 0.45391085743904114
step: 280, loss: 0.04349447414278984
step: 290, loss: 0.08113398402929306
step: 300, loss: 0.17757612466812134
step: 310, loss: 0.05589057132601738
step: 320, loss: 0.11558621376752853
step: 330, loss: 0.08534763753414154
step: 340, loss: 0.3165716528892517
step: 350, loss: 0.16055017709732056
```

```
step: 360, loss: 0.05889415740966797
step: 370, loss: 0.06767424196004868
step: 380, loss: 0.19088873267173767
step: 390, loss: 0.07158780097961426
step: 400, loss: 0.0741344764828682
step: 410, loss: 0.34556224942207336
step: 420, loss: 0.011860901489853859
step: 430, loss: 0.03327515721321106
step: 440, loss: 0.13996627926826477
step: 450, loss: 0.04138410836458206
step: 460, loss: 0.022757986560463905
step: 470, loss: 0.05310961604118347
step: 480, loss: 0.13828930258750916
step: 490, loss: 0.13869421184062958
step: 500, loss: 0.04025708884000778
step: 510, loss: 0.06413526087999344
step: 520, loss: 0.07252749055624008
step: 530, loss: 0.005756427068263292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9310504396112911, f1=0.933579335793358,
best_f1=0.933579335793358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12305629253387451
step: 10, loss: 0.08388069272041321
step: 20, loss: 0.02975926734507084
step: 30, loss: 0.049234941601753235
step: 40, loss: 0.09466506540775299
step: 50, loss: 0.060012876987457275
step: 60, loss: 0.06507151573896408
step: 70, loss: 0.01962447166442871
step: 80, loss: 0.00878225825726986
step: 90, loss: 0.01550772599875927
step: 100, loss: 0.1449766755104065
step: 110, loss: 0.004757774528115988
step: 120, loss: 0.16298280656337738
step: 130, loss: 0.01573626883327961
step: 140, loss: 0.01952546089887619
step: 150, loss: 0.05908464267849922
step: 160, loss: 0.01966383121907711
step: 170, loss: 0.03228352963924408
```

```
step: 180, loss: 0.05725819617509842
step: 190, loss: 0.006165779195725918
step: 200, loss: 0.21001118421554565
step: 210, loss: 0.05002313852310181
step: 220, loss: 0.002169228158891201
step: 230, loss: 0.05441831424832344
step: 240, loss: 0.13366280496120453
step: 250, loss: 0.011470610275864601
step: 260, loss: 0.05527602881193161
step: 270, loss: 0.10270500928163528
step: 280, loss: 0.06008220463991165
step: 290, loss: 0.06472251564264297
step: 300, loss: 0.053200021386146545
step: 310, loss: 0.027155961841344833
step: 320, loss: 0.07664640247821808
step: 330, loss: 0.06843483448028564
step: 340, loss: 0.04839438572525978
step: 350, loss: 0.0018773444462567568
step: 360, loss: 0.02843349613249302
step: 370, loss: 0.0036224147770553827
step: 380, loss: 0.18084488809108734
step: 390, loss: 0.00470300717279315
step: 400, loss: 0.12256374955177307
step: 410, loss: 0.02361072227358818
step: 420, loss: 0.12485472857952118
step: 430, loss: 0.21463553607463837
step: 440, loss: 0.028858713805675507
step: 450, loss: 0.06535506248474121
step: 460, loss: 0.0747695192694664
step: 470, loss: 0.03532596305012703
step: 480, loss: 0.008045675233006477
step: 490, loss: 0.08912461996078491
step: 500, loss: 0.008637825027108192
step: 510, loss: 0.025857385247945786
step: 520, loss: 0.2894958257675171
step: 530, loss: 0.057422176003456116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.940009447331129, f1=0.9417249417249418,
best_f1=0.9417249417249418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.07985968142747879
step: 10, loss: 0.03511710837483406
step: 20, loss: 0.030413616448640823
step: 30, loss: 0.0204522293061018
step: 40, loss: 0.02513117343187332
step: 50, loss: 0.030519232153892517
step: 60, loss: 0.03440319001674652
step: 70, loss: 0.01658038981258869
step: 80, loss: 0.05584266409277916
step: 90, loss: 0.022543029859662056
step: 100, loss: 0.008707018569111824
step: 110, loss: 0.008376878686249256
step: 120, loss: 0.12247547507286072
step: 130, loss: 0.03805556148290634
step: 140, loss: 0.01971266232430935
step: 150, loss: 0.12230619788169861
step: 160, loss: 0.08442271500825882
step: 170, loss: 0.008197986520826817
step: 180, loss: 0.08317387104034424
step: 190, loss: 0.0031657544896006584
step: 200, loss: 0.039299849420785904
step: 210, loss: 0.038734979927539825
step: 220, loss: 0.058987874537706375
step: 230, loss: 0.01759718731045723
step: 240, loss: 0.036360837519168854
step: 250, loss: 0.11009494215250015
step: 260, loss: 0.05569400265812874
step: 270, loss: 0.015754083171486855
step: 280, loss: 0.003692558268085122
step: 290, loss: 0.006842480506747961
step: 300, loss: 0.16516953706741333
step: 310, loss: 0.06276939809322357
step: 320, loss: 0.008097540587186813
step: 330, loss: 0.003123043803498149
step: 340, loss: 0.020201778039336205
step: 350, loss: 0.2623293697834015
step: 360, loss: 0.025422917678952217
step: 370, loss: 0.04309714213013649
step: 380, loss: 0.004897307604551315
step: 390, loss: 0.004620210267603397
step: 400, loss: 0.12374253571033478
step: 410, loss: 0.08365479856729507
step: 420, loss: 0.0736553817987442
step: 430, loss: 0.05308382958173752
step: 440, loss: 0.16766171157360077
step: 450, loss: 0.020746778696775436
step: 460, loss: 0.054491911083459854
step: 470, loss: 0.012603172101080418
```

```
step: 480, loss: 0.0956219807267189
step: 490, loss: 0.026353636756539345
step: 500, loss: 0.004271041136234999
step: 510, loss: 0.039052754640579224
step: 520, loss: 0.007326139137148857
step: 530, loss: 0.013833357952535152
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.940354147250699, f1=0.9319664492078286,
best_f1=0.9319664492078286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007012298796325922
step: 10, loss: 0.0006168937543407083
step: 20, loss: 0.03369581699371338
step: 30, loss: 0.0367802269756794
step: 40, loss: 0.025644036009907722
step: 50, loss: 0.0652385801076889
step: 60, loss: 0.004847624339163303
step: 70, loss: 0.01035055797547102
step: 80, loss: 0.023576725274324417
step: 90, loss: 0.03576295077800751
step: 100, loss: 0.01908167265355587
step: 110, loss: 0.1837099939584732
step: 120, loss: 0.002393315313383937
step: 130, loss: 0.10205834358930588
step: 140, loss: 0.05765815079212189
step: 150, loss: 0.05444016307592392
step: 160, loss: 0.01235256902873516
step: 170, loss: 0.03001713939011097
step: 180, loss: 0.11026198416948318
step: 190, loss: 0.048326071351766586
step: 200, loss: 0.05959578976035118
step: 210, loss: 0.03389362245798111
step: 220, loss: 0.01773049309849739
step: 230, loss: 0.00758357485756278
step: 240, loss: 0.0166025348007679
step: 250, loss: 0.04926964268088341
step: 260, loss: 0.001900924602523446
step: 270, loss: 0.07136289775371552
step: 280, loss: 0.01957153156399727
step: 290, loss: 0.010212480090558529
```

```
step: 300, loss: 0.000404133927077055
step: 310, loss: 0.03308697044849396
step: 320, loss: 0.08736483007669449
step: 330, loss: 0.018767017871141434
step: 340, loss: 0.016558414325118065
step: 350, loss: 0.024626923725008965
step: 360, loss: 0.015073433518409729
step: 370, loss: 0.01301803532987833
step: 380, loss: 0.010030830278992653
step: 390, loss: 9.337341180071235e-05
step: 400, loss: 0.0008375451434403658
step: 410, loss: 0.0007527595153078437
step: 420, loss: 0.023395130410790443
step: 430, loss: 0.0036480568815022707
step: 440, loss: 0.0019857280422002077
step: 450, loss: 0.05074739828705788
step: 460, loss: 0.12587490677833557
step: 470, loss: 0.003375882050022483
step: 480, loss: 0.20609721541404724
step: 490, loss: 0.05081714689731598
step: 500, loss: 0.044341377913951874
step: 510, loss: 0.051111526787281036
step: 520, loss: 0.0070382170379161835
step: 530, loss: 0.13854379951953888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9432029795158287, f1=0.940354147250699,
best_f1=0.940354147250699
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008957559475675225
step: 10, loss: 0.0019327541813254356
step: 20, loss: 0.003860567696392536
step: 30, loss: 0.0014755175216123462
step: 40, loss: 0.0010315874824300408
step: 50, loss: 0.13005505502223969
step: 60, loss: 0.021580658853054047
step: 70, loss: 0.021487079560756683
step: 80, loss: 0.000447409984190017
step: 90, loss: 0.00092666334239766
step: 100, loss: 0.07181833684444427
step: 110, loss: 0.0033600491005927324
```

```
step: 120, loss: 0.18439370393753052
step: 130, loss: 0.011073592118918896
step: 140, loss: 0.01968294195830822
step: 150, loss: 0.00659205811098218
step: 160, loss: 0.0030056778341531754
step: 170, loss: 0.06089628115296364
step: 180, loss: 0.0034383495803922415
step: 190, loss: 0.007707271724939346
step: 200, loss: 0.009731687605381012
step: 210, loss: 0.054489947855472565
step: 220, loss: 0.0035137240774929523
step: 230, loss: 0.09079378098249435
step: 240, loss: 0.018208084627985954
step: 250, loss: 0.2367325872182846
step: 260, loss: 0.008447430096566677
step: 270, loss: 0.0021592164412140846
step: 280, loss: 0.0022079995833337307
step: 290, loss: 0.001395181636326015
step: 300, loss: 0.008006338961422443
step: 310, loss: 0.01192989107221365
step: 320, loss: 0.011401350609958172
step: 330, loss: 0.0016230040928348899
step: 340, loss: 0.011741919443011284
step: 350, loss: 0.0035569644533097744
step: 360, loss: 0.002291072392836213
step: 370, loss: 0.0005385593976825476
step: 380, loss: 0.0003939152229577303
step: 390, loss: 0.001585142919793725
step: 400, loss: 0.0027665880043059587
step: 410, loss: 0.006676750723272562
step: 420, loss: 0.17397353053092957
step: 430, loss: 0.004134945571422577
step: 440, loss: 0.0026042030658572912
step: 450, loss: 0.04246153309941292
step: 460, loss: 0.3652077615261078
step: 470, loss: 0.06917495280504227
step: 480, loss: 0.09312575310468674
step: 490, loss: 0.007291873916983604
step: 500, loss: 0.01103593036532402
step: 510, loss: 0.0016760803991928697
step: 520, loss: 0.11646231263875961
step: 530, loss: 0.026295233517885208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9475150952159777, f1=0.9462465245597776,
best_f1=0.9462465245597776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021959353238344193
step: 10, loss: 0.00030775624327361584
step: 20, loss: 0.04150415584445
step: 30, loss: 0.0005545670283026993
step: 40, loss: 0.0008187069906853139
step: 50, loss: 0.0011747278040274978
step: 60, loss: 0.008393172174692154
step: 70, loss: 0.002333215204998851
step: 80, loss: 0.0009497018763795495
step: 90, loss: 0.004292096942663193
step: 100, loss: 0.016792168840765953
step: 110, loss: 0.0007500657229684293
step: 120, loss: 0.009130606427788734
step: 130, loss: 0.00428790831938386
step: 140, loss: 0.001467605703510344
step: 150, loss: 0.01689077541232109
step: 160, loss: 0.07322248816490173
step: 170, loss: 0.011367866769433022
step: 180, loss: 0.0019699151162058115
step: 190, loss: 0.28875285387039185
step: 200, loss: 0.010148352943360806
step: 210, loss: 0.00298207881860435
step: 220, loss: 0.009317625313997269
step: 230, loss: 0.010977608151733875
step: 240, loss: 0.019405178725719452
step: 250, loss: 0.14060784876346588
step: 260, loss: 0.006593993399292231
step: 270, loss: 0.0028813956305384636
step: 280, loss: 0.05837082117795944
step: 290, loss: 0.0006624290253967047
step: 300, loss: 0.03917786106467247
step: 310, loss: 0.04186486825346947
step: 320, loss: 7.051845022942871e-05
step: 330, loss: 0.019150346517562866
step: 340, loss: 0.0003565533261280507
step: 350, loss: 0.0011494463542476296
step: 360, loss: 0.12272074818611145
step: 370, loss: 0.0018086719792336226
step: 380, loss: 0.00040639497456140816
step: 390, loss: 0.0023624980822205544
step: 400, loss: 0.0017837153282016516
step: 410, loss: 0.001377549604512751
```

```
step: 420, loss: 0.10560908913612366
step: 430, loss: 0.002300625666975975
step: 440, loss: 0.0023659816943109035
step: 450, loss: 0.17840448021888733
step: 460, loss: 0.0038398420438170433
step: 470, loss: 0.0007886184030212462
step: 480, loss: 0.03465866670012474
step: 490, loss: 0.008161385543644428
step: 500, loss: 0.004877311643213034
step: 510, loss: 0.170180082321167
step: 520, loss: 0.0026295650750398636
step: 530, loss: 0.026255186647176743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9483960948396094, f1=0.9418334108887855,
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014554793015122414
step: 10, loss: 0.0007329101790674031
step: 20, loss: 0.0020742418710142374
step: 30, loss: 0.006486412137746811
step: 40, loss: 0.02435934729874134
step: 50, loss: 0.0009612789144739509
step: 60, loss: 0.0003323464188724756
step: 70, loss: 9.6032781584654e-05
step: 80, loss: 8.836883353069425e-05
step: 90, loss: 0.001773406402207911
step: 100, loss: 0.024122413247823715
step: 110, loss: 0.0013798072468489408
step: 120, loss: 0.002483765361830592
step: 130, loss: 0.003996985033154488
step: 140, loss: 0.00013592938194051385
step: 150, loss: 0.0002449257008265704
step: 160, loss: 0.0004697997064795345
step: 170, loss: 0.006347543094307184
step: 180, loss: 0.05912751331925392
step: 190, loss: 0.0019216843647882342
step: 200, loss: 0.00010480216587893665
step: 210, loss: 7.249864574987441e-05
step: 220, loss: 9.232413867721334e-05
step: 230, loss: 0.0005457473453134298
```

```
step: 240, loss: 0.0010844581993296742
step: 250, loss: 0.018172601237893105
step: 260, loss: 0.01556358952075243
step: 270, loss: 0.0010424550855532289
step: 280, loss: 0.018360400572419167
step: 290, loss: 0.0014838924398645759
step: 300, loss: 0.0013974191388115287
step: 310, loss: 0.0024117939174175262
step: 320, loss: 0.0022101812064647675
step: 330, loss: 0.04018373787403107
step: 340, loss: 0.001412220997735858
step: 350, loss: 0.0059793610125780106
step: 360, loss: 0.020166493952274323
step: 370, loss: 0.0022089441772550344
step: 380, loss: 0.006259097717702389
step: 390, loss: 0.006799283437430859
step: 400, loss: 0.009323619306087494
step: 410, loss: 0.00019922609499190003
step: 420, loss: 0.0015708954306319356
step: 430, loss: 0.0006965866195969284
step: 440, loss: 0.0018486082553863525
step: 450, loss: 0.0011154927778989077
step: 460, loss: 0.0018295395420864224
step: 470, loss: 0.035654935985803604
step: 480, loss: 0.0012678890489041805
step: 490, loss: 0.00364995957352221
step: 500, loss: 0.0028390998486429453
step: 510, loss: 0.0002464273711666465
step: 520, loss: 0.00013905471132602543
step: 530, loss: 0.004167531151324511
/usr/local/lib/python3.7/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=0.941340782122905, f1=0.9381107491856677,
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035973486956208944
step: 10, loss: 0.0031594918109476566
step: 20, loss: 0.0002677918237168342
step: 30, loss: 0.001660175621509552
step: 40, loss: 0.0003029982908628881
step: 50, loss: 0.0007088928250595927
step: 60, loss: 0.0023220248986035585
step: 70, loss: 0.005740913562476635
step: 80, loss: 0.0004749953805003315
step: 90, loss: 0.00015463038289453834
```

```
step: 100, loss: 0.0003640997165348381
step: 110, loss: 0.00017353967996314168
step: 120, loss: 0.0001198091558762826
step: 130, loss: 0.00195224245544523
step: 140, loss: 0.00016651830810587853
step: 150, loss: 0.00013059156481176615
step: 160, loss: 7.460382767021656e-05
step: 170, loss: 0.002288586925715208
step: 180, loss: 0.04822293296456337
step: 190, loss: 0.0019063586369156837
step: 200, loss: 0.007648531813174486
step: 210, loss: 0.05708477273583412
step: 220, loss: 0.004269807133823633
step: 230, loss: 0.019214171916246414
step: 240, loss: 0.003223974956199527
step: 250, loss: 5.108491677674465e-05
step: 260, loss: 0.00022539476049132645
step: 270, loss: 0.0010549507569521666
step: 280, loss: 6.548118108185008e-05
step: 290, loss: 6.59787910990417e-05
step: 300, loss: 0.00031254434725269675
step: 310, loss: 0.2183813750743866
step: 320, loss: 0.0002587187918834388
step: 330, loss: 0.05579583719372749
step: 340, loss: 0.023214299231767654
step: 350, loss: 0.000279231317108497
step: 360, loss: 0.017229560762643814
step: 370, loss: 0.0478602796792984
step: 380, loss: 0.00012133768905187026
step: 390, loss: 0.0007067812257446349
step: 400, loss: 0.0011112960055470467
step: 410, loss: 0.0007341663585975766
step: 420, loss: 0.008791684173047543
step: 430, loss: 0.00235192384570837
step: 440, loss: 0.050231609493494034
step: 450, loss: 0.0030817193910479546
step: 460, loss: 0.001115836901590228
step: 470, loss: 0.07494613528251648
step: 480, loss: 0.0037008372601121664
step: 490, loss: 0.0006839919369667768
step: 500, loss: 0.0003387216420378536
step: 510, loss: 0.019080406054854393
step: 520, loss: 0.001018220093101263
step: 530, loss: 0.008573259226977825
/usr/local/lib/python3.7/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=0.9418386491557222, f1=0.9380614657210402,
```

```
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011441040987847373
step: 10, loss: 0.0008026307914406061
step: 20, loss: 0.00042119371937587857
step: 30, loss: 0.069375857710838332
step: 40, loss: 0.0008677522419020534
step: 50, loss: 0.0002499435795471072
step: 60, loss: 0.00018255818577017635
step: 70, loss: 0.016280194744467735
step: 80, loss: 0.0021691613364964724
step: 90, loss: 0.22352659702301025
step: 100, loss: 0.0006200565258041024
step: 110, loss: 0.0071100347429513931
step: 120, loss: 0.0006974351126700064
step: 130, loss: 0.001147479866631329
step: 140, loss: 0.0005202900501899421
step: 150, loss: 0.016954340040683746
step: 160, loss: 0.000895594828762114
step: 170, loss: 0.000824205984827131
step: 180, loss: 0.000648347195237875
step: 190, loss: 9.789977775653824e-05
step: 200, loss: 0.002982360776513815
step: 210, loss: 0.0002798668574541807
step: 220, loss: 0.0010712533257901669
step: 230, loss: 0.0017124194419011474
step: 240, loss: 0.00015405757585540414
step: 250, loss: 0.0003362598654348403
step: 260, loss: 0.037324827164411545
step: 270, loss: 0.012123098596930504
step: 280, loss: 0.0054612308740615845
step: 290, loss: 0.00014985958114266396
step: 300, loss: 0.0002600857987999916
step: 310, loss: 0.009120318107306957
step: 320, loss: 9.231305011780933e-05
step: 330, loss: 0.00026417296612635255
step: 340, loss: 0.00029920059023424983
step: 350, loss: 0.024136709049344063
step: 360, loss: 0.0001381327019771561
step: 370, loss: 0.0002954927913378924
step: 380, loss: 0.0009657046175561845
step: 390, loss: 0.00011915928917005658
step: 400, loss: 0.08962256461381912
step: 410, loss: 0.0039978246204555035
step: 420, loss: 0.00021083236788399518
step: 430, loss: 0.01785506308078766
```

```
step: 440, loss: 0.00021224698866717517
step: 450, loss: 0.0041630566120147705
step: 460, loss: 0.0001988807925954461
step: 470, loss: 9.904029866447672e-05
step: 480, loss: 0.007761497050523758
step: 490, loss: 0.0012898111017420888
step: 500, loss: 0.0002977332624141127
step: 510, loss: 0.12014591693878174
step: 520, loss: 0.0033582081086933613
step: 530, loss: 0.004384715110063553
/usr/local/lib/python3.7/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=0.9475638051044084, f1=0.9441860465116279,
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012541133037302643
step: 10, loss: 0.0007088237907737494
step: 20, loss: 8.841138333082199e-05
step: 30, loss: 0.0010932041332125664
step: 40, loss: 0.00019659186364151537
step: 50, loss: 0.00040547276148572564
step: 60, loss: 0.00016447989037260413
step: 70, loss: 6.217951158760116e-05
step: 80, loss: 0.0034322189167141914
step: 90, loss: 0.0004433783469721675
step: 100, loss: 0.0008186919149011374
step: 110, loss: 0.00737516675144434
step: 120, loss: 0.0001411479024682194
step: 130, loss: 6.675897748209536e-05
step: 140, loss: 0.0001301837182836607
step: 150, loss: 0.0008236707071773708
step: 160, loss: 0.0019183441763743758
step: 170, loss: 0.00026075300411321223
step: 180, loss: 0.0010387359652668238
step: 190, loss: 3.695100167533383e-05
step: 200, loss: 0.00027931114891543984
step: 210, loss: 0.0037016337737441063
step: 220, loss: 0.00013971547014079988
step: 230, loss: 0.019570186734199524
step: 240, loss: 0.000115383074444253
step: 250, loss: 0.00012033770326524973
step: 260, loss: 0.0001445917587261647
step: 270, loss: 0.0014069888275116682
step: 280, loss: 0.00442067626863718
step: 290, loss: 9.568949462845922e-05
```

```
step: 300, loss: 0.004360207822173834
step: 310, loss: 0.05583760142326355
step: 320, loss: 0.025731531903147697
step: 330, loss: 5.594022513832897e-05
step: 340, loss: 3.505656059132889e-05
step: 350, loss: 0.000596099067479372
step: 360, loss: 0.00066369742853567
step: 370, loss: 7.408962119370699e-05
step: 380, loss: 5.382924064178951e-05
step: 390, loss: 4.074004027643241e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 400, loss: 4.685382737079635e-05
step: 410, loss: 0.0030094031244516373
step: 420, loss: 6.99907832313329e-05
step: 430, loss: 5.0342110625933856e-05
step: 440, loss: 2.80876975011779e-05
step: 450, loss: 0.0008367178961634636
step: 460, loss: 0.0010553330648690462
step: 470, loss: 0.001361255650408566
step: 480, loss: 0.00044692904339171946
step: 490, loss: 0.03274356946349144
step: 500, loss: 0.008871848694980145
step: 510, loss: 2.1635894881910644e-05
step: 520, loss: 0.000278961262665689
step: 530, loss: 8.800683281151578e-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=0.9471715755025713, f1=0.943502824858757,
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.819569817802403e-05
step: 10, loss: 0.0006358372629620135
step: 20, loss: 4.957098644808866e-05
step: 30, loss: 6.506164208985865e-05
step: 40, loss: 0.0004079641366843134
step: 50, loss: 2.4969731384771876e-05
step: 60, loss: 0.00014885971904732287
step: 70, loss: 4.6777418901911005e-05
step: 80, loss: 0.04164653643965721
step: 90, loss: 0.007889318279922009
step: 100, loss: 0.00013756340194959193
step: 110, loss: 0.0003539726894814521
step: 120, loss: 0.017861705273389816
step: 130, loss: 0.00015107335639186203
step: 140, loss: 0.0001826749212341383
```

```
step: 150, loss: 0.00013129755097907037
step: 160, loss: 0.0007214279612526298
step: 170, loss: 0.0007403569761663675
step: 180, loss: 0.0004157565417699516
step: 190, loss: 0.0037017501890659332
step: 200, loss: 0.00023383561347145587
step: 210, loss: 0.0002558564592618495
step: 220, loss: 0.007277178578078747
step: 230, loss: 2.5535018721711822e-05
step: 240, loss: 0.0014251737156882882
step: 250, loss: 0.0011149125639349222
step: 260, loss: 7.00490054441616e-05
step: 270, loss: 0.0001804429484764114
step: 280, loss: 0.00042141074663959444
step: 290, loss: 0.0001846997329266742
step: 300, loss: 4.580391032504849e-05
step: 310, loss: 6.944852066226304e-05
step: 320, loss: 2.1084613763378002e-05
step: 330, loss: 1.978071486519184e-05
step: 340, loss: 0.0036205467768013477
step: 350, loss: 0.01967841573059559
step: 360, loss: 6.723216210957617e-05
step: 370, loss: 2.024638160946779e-05
step: 380, loss: 0.00026321300538256764
step: 390, loss: 0.00010915938764810562
step: 400, loss: 7.58782189222984e-05
step: 410, loss: 7.622831617482007e-05
step: 420, loss: 0.00036927350447513163
step: 430, loss: 0.00033476867247372866
step: 440, loss: 0.0006034895777702332
step: 450, loss: 0.015986187383532524
step: 460, loss: 0.00015062079182825983
step: 470, loss: 0.00021189880499150604
step: 480, loss: 0.0004695398092735559
step: 490, loss: 0.007721039932221174
step: 500, loss: 0.0001659808331169188
step: 510, loss: 3.8557351217605174e-05
step: 520, loss: 0.0006386323366314173
step: 530, loss: 0.15043002367019653
/usr/local/lib/python3.7/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=0.9471733086190918, f1=0.9460083064143978,
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.924323290353641e-05
```

```
step: 10, loss: 0.002435389906167984
step: 20, loss: 0.0034951416309922934
step: 30, loss: 5.942126881564036e-05
step: 40, loss: 0.0004848049720749259
step: 50, loss: 0.00024370974279008806
step: 60, loss: 4.401158366817981e-05
step: 70, loss: 5.3239935368765146e-05
step: 80, loss: 0.0003052822139579803
step: 90, loss: 0.00011972539505222812
step: 100, loss: 0.028872592374682426
step: 110, loss: 0.0032291882671415806
step: 120, loss: 0.0006881621666252613
step: 130, loss: 0.0003009644860867411
step: 140, loss: 0.00033729089773260057
step: 150, loss: 0.00028840589220635593
step: 160, loss: 0.00022478536993730813
step: 170, loss: 0.00011487772280815989
step: 180, loss: 0.0001526198029750958
step: 190, loss: 0.0008215407142415643
step: 200, loss: 0.008806224912405014
step: 210, loss: 0.0007023524958640337
step: 220, loss: 7.106227712938562e-05
step: 230, loss: 0.001731419120915234
step: 240, loss: 8.819958748063073e-05
step: 250, loss: 0.002000223146751523
step: 260, loss: 8.202054596040398e-05
step: 270, loss: 0.0008522222051396966
step: 280, loss: 0.0006024212343618274
step: 290, loss: 0.0003996889863628894
step: 300, loss: 0.000292181532131508
step: 310, loss: 0.0005075421649962664
step: 320, loss: 0.00011681944306474179
step: 330, loss: 0.00013683237193617977
step: 340, loss: 0.0059562972746789455
step: 350, loss: 3.2841984648257494e-05
step: 360, loss: 0.0017350605921819806
step: 370, loss: 6.165004015201703e-05
step: 380, loss: 0.00011312399874441326
step: 390, loss: 0.02530919387936592
step: 400, loss: 0.00024633383145555854
step: 410, loss: 8.34328675409779e-05
step: 420, loss: 7.700815331190825e-05
step: 430, loss: 0.00017142071737907827
step: 440, loss: 0.0003843736194539815
step: 450, loss: 0.003665247932076454
step: 460, loss: 6.572987331310287e-05
step: 470, loss: 0.00022594985784962773
step: 480, loss: 0.00028691161423921585
```

```
step: 490, loss: 6.076179488445632e-05
step: 500, loss: 1.864087062131148e-05
step: 510, loss: 0.00016326192417182028
step: 520, loss: 0.00013719081471208483
step: 530, loss: 6.906808994244784e-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=0.9450961989676209, f1=0.9400749063670412,
best_f1=0.9418334108887855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.176335616852157e-05
step: 10, loss: 7.212423224700615e-05
step: 20, loss: 4.982881000614725e-05
step: 30, loss: 2.0112354832235724e-05
step: 40, loss: 0.000654426054097712
step: 50, loss: 0.0012279421789571643
step: 60, loss: 0.00011778737825807184
step: 70, loss: 0.0006358959362842143
step: 80, loss: 0.00031363443122224537
step: 90, loss: 5.015612987335771e-05
step: 100, loss: 6.446550833061337e-05
step: 110, loss: 0.00012453942326828837
step: 120, loss: 5.127099939272739e-05
step: 130, loss: 0.00011385418474674225
step: 140, loss: 0.00031582586234433544
step: 150, loss: 0.00014952574565541
step: 160, loss: 0.0001596535148564726
step: 170, loss: 0.006063214503228664
step: 180, loss: 3.491121970000677e-05
step: 190, loss: 0.00014045096759218723
step: 200, loss: 6.768291495973244e-05
step: 210, loss: 0.00017846225819084793
step: 220, loss: 0.00011272755364188924
step: 230, loss: 0.01333710178732872
step: 240, loss: 7.58706228225492e-05
step: 250, loss: 0.000192242136108689
step: 260, loss: 2.9215765607659705e-05
step: 270, loss: 0.00011659042502287775
step: 280, loss: 0.00019640295067802072
step: 290, loss: 2.9266480851219967e-05
step: 300, loss: 1.7538328393129632e-05
step: 310, loss: 0.00010214224312221631
step: 320, loss: 2.5725083105498925e-05
step: 330, loss: 5.7379216741537675e-05
step: 340, loss: 9.32335969991982e-05
```

```
step: 350, loss: 1.0430698239360936e-05
step: 360, loss: 0.001878432696685195
step: 370, loss: 0.00010457854659762233
step: 380, loss: 0.001115289400331676
step: 390, loss: 2.913359821832273e-05
step: 400, loss: 0.00014823544188402593
step: 410, loss: 4.2651372496038675e-05
step: 420, loss: 0.0029473500326275826
step: 430, loss: 5.680361209670082e-05
step: 440, loss: 2.9023825845797546e-05
step: 450, loss: 6.852936348877847e-05
step: 460, loss: 0.00017876231868285686
step: 470, loss: 1.1946763152081985e-05
step: 480, loss: 8.448879270872567e-06
step: 490, loss: 2.4667635443620384e-05
step: 500, loss: 1.0523802302486729e-05
step: 510, loss: 4.168771192780696e-05
step: 520, loss: 4.5539276470663026e-05
step: 530, loss: 4.9569982365937904e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9495136637332099, f1=0.9448818897637795,
best_f1=0.9448818897637795
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.5173753935378045e-05
step: 10, loss: 0.0020014343317598104
step: 20, loss: 2.0037407011841424e-05
step: 30, loss: 0.008381200022995472
step: 40, loss: 1.3842942280462012e-05
step: 50, loss: 2.797873275994789e-05
step: 60, loss: 0.0002832739846780896
step: 70, loss: 0.0007863297942094505
step: 80, loss: 0.0001562020042911172
step: 90, loss: 1.591771979292389e-05
step: 100, loss: 1.5828471077838913e-05
step: 110, loss: 4.8593676183372736e-05
step: 120, loss: 1.6390964447055012e-05
step: 130, loss: 3.170175477862358e-05
step: 140, loss: 2.0931938706780784e-05
step: 150, loss: 2.2410278688766994e-05
step: 160, loss: 3.485343404463492e-05
```

```
step: 170, loss: 8.272183185908943e-05
step: 180, loss: 0.0027633383870124817
step: 190, loss: 0.000124953527119942
step: 200, loss: 4.3623786041280255e-05
step: 210, loss: 5.0747439672704786e-05
step: 220, loss: 1.2583845091285184e-05
step: 230, loss: 0.00010862015915336087
step: 240, loss: 1.1164403076691087e-05
step: 250, loss: 2.69423962890869e-05
step: 260, loss: 0.00011520038970047608
step: 270, loss: 0.0002810654987115413
step: 280, loss: 3.354158252477646e-05
step: 290, loss: 1.3987959391670302e-05
step: 300, loss: 1.778413206920959e-05
step: 310, loss: 2.938171019195579e-05
step: 320, loss: 0.0005974984960630536
step: 330, loss: 2.03676063392777e-05
step: 340, loss: 4.2902393033728004e-05
step: 350, loss: 1.3042034879617859e-05
step: 360, loss: 0.009535940364003181
step: 370, loss: 2.253934508189559e-05
step: 380, loss: 0.0018056617118418217
step: 390, loss: 2.669143032107968e-05
step: 400, loss: 0.00010966135596390814
step: 410, loss: 8.936875929066446e-06
step: 420, loss: 1.310162406298332e-05
step: 430, loss: 9.851121285464615e-05
step: 440, loss: 5.250032336334698e-05
step: 450, loss: 2.149354259017855e-05
step: 460, loss: 0.020049545913934708
step: 470, loss: 1.7251482859137468e-05
step: 480, loss: 1.778755904524587e-05
step: 490, loss: 0.0016873243730515242
step: 500, loss: 0.00012838139082305133
step: 510, loss: 6.984356878092512e-05
step: 520, loss: 1.028910992317833e-05
step: 530, loss: 1.992588113353122e-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=0.9497206703910613, f1=0.9448082319925164,
best_f1=0.9448082319925164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.134224671065062e-05
step: 10, loss: 1.8808559616445564e-05
step: 20, loss: 0.00015062409511301666
step: 30, loss: 0.0001502198720118031
step: 40, loss: 1.4137281141302083e-05
step: 50, loss: 0.020468614995479584
step: 60, loss: 4.4425287342164665e-05
step: 70, loss: 1.3768381904810667e-05
step: 80, loss: 1.84097279998241e-05
step: 90, loss: 1.4025470591150224e-05
step: 100, loss: 1.3153477993910201e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.0008183519239537418
step: 120, loss: 1.1321007150399964e-05
step: 130, loss: 1.3768461940344423e-05
step: 140, loss: 7.033252768451348e-05
step: 150, loss: 2.6086889192811213e-05
step: 160, loss: 1.8703223759075627e-05
step: 170, loss: 3.622725125751458e-05
step: 180, loss: 1.7106207451433875e-05
step: 190, loss: 1.149606760009192e-05
step: 200, loss: 1.550376691739075e-05
step: 210, loss: 1.1544480912561994e-05
step: 220, loss: 4.632324635167606e-05
step: 230, loss: 8.277631422970444e-05
step: 240, loss: 1.3377285540627781e-05
step: 250, loss: 1.2777449228451587e-05
step: 260, loss: 1.0620659850246739e-05
step: 270, loss: 2.5344894311274402e-05
step: 280, loss: 1.1063968486269005e-05
step: 290, loss: 3.084992931690067e-05
step: 300, loss: 1.73892367456574e-05
step: 310, loss: 1.8957423890242353e-05
step: 320, loss: 2.4444259906886145e-05
step: 330, loss: 5.5209417041623965e-05
step: 340, loss: 1.8015112800640054e-05
step: 350, loss: 1.5075916962814517e-05
step: 360, loss: 2.9062735848128796e-05
step: 370, loss: 1.8044949683826417e-05
step: 380, loss: 1.9083898223470896e-05
step: 390, loss: 1.8163915228797123e-05
step: 400, loss: 0.02535376511514187
step: 410, loss: 2.3396289179800078e-05
step: 420, loss: 1.4315864063974004e-05
step: 430, loss: 1.1622528290899936e-05
step: 440, loss: 1.2460925972845871e-05
step: 450, loss: 2.530130404920783e-05
```

```
step: 460, loss: 1.885262827272527e-05
step: 470, loss: 1.0937311344605405e-05
step: 480, loss: 0.0007480218773707747
step: 490, loss: 0.00010825679783010855
step: 500, loss: 3.0541476007783785e-05
step: 510, loss: 1.0248144462821074e-05
step: 520, loss: 1.7411401131539606e-05
step: 530, loss: 1.214050189446425e-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=0.9496738117427772, f1=0.9454036397573495,
best_f1=0.9448082319925164
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, 246.46it/s]
load_f1 = 0.9474174034434621
real_f1 = 0.9455560725919032
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, 195.45it/s]
```

4 DITTO

4.1 DITTO STRUCTURED

4.1.1 Beer - Running the matcher

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

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

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

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

359

```
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: 0.4483066201210022
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216: UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.32941176470588235, f1=0.29545454545454547, best_f1=0.29545454545454547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42730480432510376

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.44897959183673464, f1=0.32786885245901637,
best_f1=0.32786885245901637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3994828462600708
/usr/local/lib/python3.7/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=0.3939393939393939, f1=0.35135135135135137,
best_f1=0.32786885245901637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2630670666694641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4615384615384615, f1=0.36363636363636365,
best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2676527500152588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.631578947368421, f1=0.46153846153846156,
best_f1=0.46153846153846156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25582051277160645
```

```
/usr/local/lib/python3.7/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=0.37837837837837834, f1=0.3823529411764706,
best_f1=0.46153846153846156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5159943103790283
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.6363636363636364, f1=0.4827586206896552,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4642553925514221
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.42553191489361697,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31867101788520813
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.38095238095238093,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27269884943962097
/usr/local/lib/python3.7/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=0.5853658536585367, f1=0.38095238095238093,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2664344608783722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7692307692307692, f1=0.5185185185185186,
best_f1=0.5185185185185186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12704063951969147
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4666666666666667,
best_f1=0.4666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12836483120918274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8148148148148148, f1=0.5517241379310344,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08226817101240158
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.5714285714285714,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12986621260643005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8148148148148148, f1=0.5714285714285714,
best_f1=0.5517241379310344
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, 125925.99it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7586206896551724
real_f1 = 0.7333333333333334
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.70it/s]
```

4.1.2 DBLP-ACM - Running the matcher

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


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

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

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

Defaults for this optimization level are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic

```
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.601044774055481
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.48320338129997253
step: 20, loss: 0.4665757715702057
step: 30, loss: 0.3342188000679016
step: 40, loss: 0.2948457896709442
step: 50, loss: 0.26523062586784363
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.18076762557029724
step: 70, loss: 0.2651118338108063
step: 80, loss: 0.0585084930062294
step: 90, loss: 0.1724904328584671
step: 100, loss: 0.10525801032781601
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 110, loss: 0.21471403539180756
step: 120, loss: 0.02925107441842556
step: 130, loss: 0.00525312963873148
step: 140, loss: 0.008485449478030205
step: 150, loss: 0.12716513872146606
step: 160, loss: 0.0037773142103105783
step: 170, loss: 0.030391337350010872
step: 180, loss: 0.05092627555131912
step: 190, loss: 0.01562885008752346
step: 200, loss: 0.018956471234560013
step: 210, loss: 0.009171651676297188
step: 220, loss: 0.1100674644112587
step: 230, loss: 0.07656567543745041
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9164926931106471, f1=0.9379014989293362,
best_f1=0.9379014989293362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014595226384699345
step: 10, loss: 0.28889745473861694
step: 20, loss: 0.08465345948934555
step: 30, loss: 0.01628636196255684
step: 40, loss: 0.01316377054899931
step: 50, loss: 0.11251402646303177
step: 60, loss: 0.005110040307044983
step: 70, loss: 0.005442473571747541
step: 80, loss: 0.004254545550793409
step: 90, loss: 0.0026613983791321516
step: 100, loss: 0.005429273471236229
step: 110, loss: 0.025795001536607742
step: 120, loss: 0.013375885784626007
step: 130, loss: 0.002377074910327792
step: 140, loss: 0.3459623157978058
step: 150, loss: 0.13053996860980988
step: 160, loss: 0.006436098832637072
step: 170, loss: 0.001086426549591124
step: 180, loss: 0.0017458363436162472
step: 190, loss: 0.07267861068248749
step: 200, loss: 0.004959431942552328
step: 210, loss: 0.14367765188217163
step: 220, loss: 0.014621727168560028
step: 230, loss: 0.023840660229325294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775280898876404, 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: 0.013775704428553581
step: 10, loss: 0.029507644474506378
step: 20, loss: 0.004268304910510778
step: 30, loss: 0.0015673708403483033
step: 40, loss: 0.012203117832541466

```
step: 50, loss: 0.038183629512786865
step: 60, loss: 0.0028590864967554808
step: 70, loss: 0.004381336737424135
step: 80, loss: 0.009051538072526455
step: 90, loss: 0.005967793054878712
step: 100, loss: 0.003544081700965762
step: 110, loss: 0.00249567162245512
step: 120, loss: 0.006631437223404646
step: 130, loss: 0.003139777109026909
step: 140, loss: 0.0004853863501921296
step: 150, loss: 0.05614246800541878
step: 160, loss: 0.01432800106704235
step: 170, loss: 0.0021178757306188345
step: 180, loss: 0.011968682520091534
step: 190, loss: 0.041168030351400375
step: 200, loss: 0.027963340282440186
step: 210, loss: 0.005641786381602287
step: 220, loss: 0.006866464391350746
step: 230, loss: 0.031031234189867973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819819819819819, f1=0.9795454545454545,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012472480535507202
step: 10, loss: 0.002707868814468384
step: 20, loss: 0.013105001300573349
step: 30, loss: 0.002194990636780858
step: 40, loss: 0.16675956547260284
step: 50, loss: 0.007815603166818619
step: 60, loss: 0.0024118139408528805
step: 70, loss: 0.005007002968341112
step: 80, loss: 0.0028606869745999575
step: 90, loss: 0.00546172633767128
step: 100, loss: 0.0023681826423853636
step: 110, loss: 0.0007316062110476196
step: 120, loss: 0.004299483262002468
step: 130, loss: 0.005430647172033787
step: 140, loss: 0.0018901531584560871
step: 150, loss: 0.026932781562200913
step: 160, loss: 0.010863137431442738
```

```
step: 170, loss: 0.001893070642836392
step: 180, loss: 0.13198213279247284
step: 190, loss: 0.0018562866607680917
step: 200, loss: 0.007760187145322561
step: 210, loss: 0.000938494224101305
step: 220, loss: 0.0006731275352649391
step: 230, loss: 0.001976122846826911
/usr/local/lib/python3.7/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=0.9799107142857142, f1=0.9831271091113611,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0193692184984684
step: 10, loss: 0.0019809738732874393
step: 20, loss: 0.0050527239218354225
step: 30, loss: 0.0008468670421279967
step: 40, loss: 0.004040813539177179
step: 50, loss: 0.001480477279983461
step: 60, loss: 0.000943611899856478
step: 70, loss: 0.01609034091234207
step: 80, loss: 0.048629287630319595
step: 90, loss: 0.007974610663950443
step: 100, loss: 0.0009245069231837988
step: 110, loss: 0.002423509955406189
step: 120, loss: 0.007844732142984867
step: 130, loss: 0.0016171999741345644
step: 140, loss: 0.018730254843831062
step: 150, loss: 0.004436053801327944
step: 160, loss: 0.0015196611639112234
step: 170, loss: 0.0033019750844687223
step: 180, loss: 0.007183928973972797
step: 190, loss: 0.0113056106492877
step: 200, loss: 0.003913853317499161
step: 210, loss: 0.006966515444219112
step: 220, loss: 0.0010284652234986424
step: 230, loss: 0.014276559464633465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843400447427293, 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: 0.0005794060998596251
step: 10, loss: 0.003126580035313964
step: 20, loss: 0.021726712584495544
step: 30, loss: 0.0006865723989903927
step: 40, loss: 0.000587341608479619
step: 50, loss: 0.005143911112099886
step: 60, loss: 0.019570492208003998
step: 70, loss: 0.00786567758768797
step: 80, loss: 0.01694647967815399
step: 90, loss: 0.020270884037017822
step: 100, loss: 0.0008578643319196999
step: 110, loss: 0.06829030811786652
step: 120, loss: 0.0015600338811054826
step: 130, loss: 0.015058585442602634
step: 140, loss: 0.0015396720264106989
step: 150, loss: 0.0019944575615227222
step: 160, loss: 0.0008267790544778109
step: 170, loss: 0.0002878379891626537
step: 180, loss: 0.001037842477671802
step: 190, loss: 0.1304633915424347
step: 200, loss: 0.05203157290816307
step: 210, loss: 0.010999171063303947
step: 220, loss: 0.012601600028574467
step: 230, loss: 0.0011058984091505408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.985539488320356, f1=0.9766925638179801,
best_f1=0.9766925638179801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023234013933688402
step: 10, loss: 0.0009855942334979773
step: 20, loss: 0.0005129579803906381
step: 30, loss: 0.000094576709670946
step: 40, loss: 0.010753321461379528
step: 50, loss: 0.0007339878939092159
step: 60, loss: 0.000668079243041575
step: 70, loss: 0.0002133097586920485
step: 80, loss: 0.000457749207271263
```

```
step: 90, loss: 0.00018777338846120983
step: 100, loss: 0.00048208728549070656
step: 110, loss: 0.0006349362665787339
step: 120, loss: 0.001179393962956965
step: 130, loss: 0.005172756500542164
step: 140, loss: 0.0004400248872116208
step: 150, loss: 0.002582872984930873
step: 160, loss: 0.001362670212984085
step: 170, loss: 0.010266621597111225
step: 180, loss: 0.0017386812251061201
step: 190, loss: 0.001751374569721152
step: 200, loss: 0.005899721756577492
step: 210, loss: 0.007704524789005518
step: 220, loss: 0.00037623022217303514
step: 230, loss: 0.015895122662186623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876265466816648, f1=0.9774774774774775,
best_f1=0.9774774774774775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007918314076960087
step: 10, loss: 0.004250246565788984
step: 20, loss: 0.0018663587979972363
step: 30, loss: 0.002040407620370388
step: 40, loss: 0.0027496819384396076
step: 50, loss: 0.005764123983681202
step: 60, loss: 0.0017244844930246472
step: 70, loss: 0.00027609552489593625
step: 80, loss: 0.000601413135882467
step: 90, loss: 0.0014999634586274624
step: 100, loss: 0.0006870444631204009
step: 110, loss: 0.000597676495090127
step: 120, loss: 0.0020542799029499292
step: 130, loss: 0.0020189268980175257
step: 140, loss: 0.001776599558070302
step: 150, loss: 0.1716025471687317
step: 160, loss: 0.010955479927361012
step: 170, loss: 0.021751653403043747
step: 180, loss: 0.0016115889884531498
step: 190, loss: 0.011096077971160412
step: 200, loss: 0.0056197429075837135
```

```
step: 210, loss: 0.0010789192747324705
step: 220, loss: 0.14555154740810394
step: 230, loss: 0.004584263078868389
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9852440408626559,
best_f1=0.9774774774774775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011541745625436306
step: 10, loss: 0.0013753409730270505
step: 20, loss: 0.012016917578876019
step: 30, loss: 0.0015758227091282606
step: 40, loss: 0.0017901912797242403
step: 50, loss: 0.0022306039463728666
step: 60, loss: 0.0010305464966222644
step: 70, loss: 0.19801712036132812
step: 80, loss: 0.0035469327121973038
step: 90, loss: 0.07786226272583008
step: 100, loss: 0.0013590999878942966
step: 110, loss: 0.0014837684575468302
step: 120, loss: 0.008620262145996094
step: 130, loss: 0.02291172556579113
step: 140, loss: 0.008449813351035118
step: 150, loss: 0.024664564058184624
step: 160, loss: 0.00808679684996605
step: 170, loss: 0.0002821534581016749
step: 180, loss: 0.0008106095483526587
step: 190, loss: 0.00019176804926246405
step: 200, loss: 0.00032598982215858996
step: 210, loss: 0.010751327499747276
step: 220, loss: 0.0001745121116982773
step: 230, loss: 0.00015144646749831736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876543209876544, f1=0.9807474518686297,
best_f1=0.9807474518686297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017271008982788771
```

```
step: 10, loss: 0.00030950785730965436
step: 20, loss: 0.01629357412457466
step: 30, loss: 0.00028670087340287864
step: 40, loss: 0.0008102741558104753
step: 50, loss: 0.005466127302497625
step: 60, loss: 0.0015417037066072226
step: 70, loss: 0.00047962323878891766
step: 80, loss: 0.00012980290921404958
step: 90, loss: 0.0006350053008645773
step: 100, loss: 0.00021831953199580312
step: 110, loss: 0.0007379525923170149
step: 120, loss: 0.00019538153719622642
step: 130, loss: 0.0016959882341325283
step: 140, loss: 0.0005582401645369828
step: 150, loss: 0.0028211884200572968
step: 160, loss: 8.126229658955708e-05
step: 170, loss: 0.0006637843325734138
step: 180, loss: 0.005307419691234827
step: 190, loss: 0.0036550944205373526
step: 200, loss: 0.00028480199398472905
step: 210, loss: 0.03667697682976723
step: 220, loss: 0.004527240060269833
step: 230, loss: 0.00010789640509756282
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9899216125419933, f1=0.9865470852017937,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.632541382918134e-05
step: 10, loss: 0.000257340056123212
step: 20, loss: 0.00017388336709700525
step: 30, loss: 0.00011616630217758939
step: 40, loss: 3.1169816793408245e-05
step: 50, loss: 5.3088326239958405e-05
step: 60, loss: 0.012571956031024456
step: 70, loss: 0.0001293433888349682
step: 80, loss: 0.0016040864866226912
step: 90, loss: 0.004099762067198753
step: 100, loss: 0.0027715060859918594
step: 110, loss: 0.0054379976354539394
step: 120, loss: 0.0012751342728734016
```

```
step: 130, loss: 0.0003984151699114591
step: 140, loss: 0.00446274084970355
step: 150, loss: 0.00018704206740949303
step: 160, loss: 0.00018490089860279113
step: 170, loss: 0.00031691318145021796
step: 180, loss: 0.00046788007603026927
step: 190, loss: 9.413088264409453e-05
step: 200, loss: 0.008986656554043293
step: 210, loss: 0.000113696776679717
step: 220, loss: 0.00021924893371760845
step: 230, loss: 0.00011232794349780306
/usr/local/lib/python3.7/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=0.9841269841269841,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005018181982450187
step: 10, loss: 7.264232408488169e-05
step: 20, loss: 0.002138944575563073
step: 30, loss: 0.025457637384533882
step: 40, loss: 0.00010496692266315222
step: 50, loss: 0.002119335811585188
step: 60, loss: 0.0001708668132778257
step: 70, loss: 0.00031780972494743764
step: 80, loss: 0.0003045475168619305
step: 90, loss: 0.000639884325210005
step: 100, loss: 9.546345245325938e-05
step: 110, loss: 0.002490903018042445
step: 120, loss: 0.0003016942064277828
step: 130, loss: 0.0001649767509661615
step: 140, loss: 0.00018742316751740873
step: 150, loss: 0.00035405048402026296
step: 160, loss: 0.0005689786048606038
step: 170, loss: 0.0004592682234942913
step: 180, loss: 0.00013730177306570113
step: 190, loss: 0.00592011446133256
step: 200, loss: 9.288136061513796e-05
step: 210, loss: 0.01744849421083927
step: 220, loss: 0.007105683442205191
step: 230, loss: 0.00024957803543657064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910313901345291, 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: 0.0007767750135235488
step: 10, loss: 0.00011877674842253327
step: 20, loss: 0.002722689649090171
step: 30, loss: 0.00011868976434925571
step: 40, loss: 0.0003055871347896755
step: 50, loss: 0.0038300217129290104
step: 60, loss: 0.0001822462654672563
step: 70, loss: 0.0002085308515233919
step: 80, loss: 0.010905509814620018
step: 90, loss: 0.00012181857891846448
step: 100, loss: 0.0020712711848318577
step: 110, loss: 0.004682373255491257
step: 120, loss: 0.0001924124953802675
step: 130, loss: 0.0002651257091201842
step: 140, loss: 8.396869816351682e-05
step: 150, loss: 9.8001430160366e-05
step: 160, loss: 0.0004174430505372584
step: 170, loss: 0.0002534719242248684
step: 180, loss: 0.021720807999372482
step: 190, loss: 0.000155858404468745
step: 200, loss: 3.833581286016852e-05
step: 210, loss: 0.0003276397765148431
step: 220, loss: 9.074514673557132e-05
step: 230, loss: 0.00014522098354063928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9910313901345291, f1=0.9831271091113611,
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: 0.00018934777472168207
step: 10, loss: 0.00014521963021252304
step: 20, loss: 0.00016004602366592735
step: 30, loss: 0.0003741731634363532
step: 40, loss: 0.000330751936417073
step: 50, loss: 7.032846770016477e-05
step: 60, loss: 0.0001308116625295952
step: 70, loss: 0.00018829747568815947
step: 80, loss: 9.11768656806089e-05
```

```
step: 90, loss: 0.0009064007317647338
step: 100, loss: 0.0003052198444493115
step: 110, loss: 0.0002605654881335795
step: 120, loss: 4.52119238616433e-05
step: 130, loss: 0.00017868992290459573
step: 140, loss: 0.0006145171355456114
step: 150, loss: 6.436803232645616e-05
step: 160, loss: 0.0002215390995843336
step: 170, loss: 0.00033810496097430587
step: 180, loss: 6.741531251464039e-05
step: 190, loss: 0.0015327362343668938
step: 200, loss: 0.00016147452697622352
step: 210, loss: 0.00045742525253444491
step: 220, loss: 0.0001462228101445362
step: 230, loss: 8.784316742094234e-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=0.9865771812080537, f1=0.9854423292273236,
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: 0.01281734462827444
step: 10, loss: 0.00014499401731882244
step: 20, loss: 0.0012351800687611103
step: 30, loss: 0.00010520635987631977
step: 40, loss: 0.0013469448313117027
step: 50, loss: 9.136625885730609e-05
step: 60, loss: 0.040407661348581314
step: 70, loss: 0.012499571777880192
step: 80, loss: 0.00010064652451546863
step: 90, loss: 0.0009453100501559675
step: 100, loss: 6.531969120260328e-05
step: 110, loss: 0.00015411134518217295
step: 120, loss: 0.053022440522909164
step: 130, loss: 0.000195447908481583
step: 140, loss: 0.007810922805219889
step: 150, loss: 0.0005883745034225285
step: 160, loss: 0.02240608073771
step: 170, loss: 6.059361476218328e-05
step: 180, loss: 0.0001327188656432554
step: 190, loss: 0.0003843264712486416
step: 200, loss: 0.00038542377296835184
step: 210, loss: 0.0012860994320362806
step: 220, loss: 0.00041751432581804693
step: 230, loss: 0.002301696687936783
/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9818594104308391,
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 [00:12, 204.79it/s]
load_f1 = 0.9898762654668166
real_f1 = 0.9876543209876544
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:23, 190.23it/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 roberta \
  --fp16 \
  --save_model \
  --dk general  \
  --dk product \
  --summarize \
  --da entry_swap  \
  --da attr_del


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

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

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

Defaults for this optimization level are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
```

```
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6167466044425964
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.3919159173965454
step: 20, loss: 0.33408084511756897
step: 30, loss: 0.367494136095047
step: 40, loss: 0.3537997901439667
step: 50, loss: 0.5522924661636353
step: 60, loss: 0.2852359712123871
step: 70, loss: 0.1843569576740265
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 80, loss: 0.2566659152507782
step: 90, loss: 0.17617177963256836
step: 100, loss: 0.513187825679779
step: 110, loss: 0.16649635136127472
step: 120, loss: 0.204820454120636
step: 130, loss: 0.17062415182590485
step: 140, loss: 0.29729315638542175
step: 150, loss: 0.04989166930317879
step: 160, loss: 0.18305492401123047
step: 170, loss: 0.05522830784320831
step: 180, loss: 0.03873808681964874
step: 190, loss: 0.06063690781593323
step: 200, loss: 0.031342312693595886
step: 210, loss: 0.07666359096765518
step: 220, loss: 0.09882497042417526
step: 230, loss: 0.2838962972164154
step: 240, loss: 0.08412319421768188
step: 250, loss: 0.09355112165212631
step: 260, loss: 0.13715283572673798
step: 270, loss: 0.25384122133255005
step: 280, loss: 0.10894577205181122
```

```
step: 290, loss: 0.04219679906964302
step: 300, loss: 0.04244177043437958
step: 310, loss: 0.25425633788108826
step: 320, loss: 0.04617707058787346
step: 330, loss: 0.055115267634391785
step: 340, loss: 0.5895943641662598
step: 350, loss: 0.1450272798538208
step: 360, loss: 0.07455772906541824
step: 370, loss: 0.007628240156918764
step: 380, loss: 0.24273136258125305
step: 390, loss: 0.029060641303658485
step: 400, loss: 0.033032409846782684
step: 410, loss: 0.21616697311401367
step: 420, loss: 0.007457866799086332
step: 430, loss: 0.12388689070940018
step: 440, loss: 0.06806082278490067
step: 450, loss: 0.0653345063328743
step: 460, loss: 0.016721144318580627
step: 470, loss: 0.03636949136853218
step: 480, loss: 0.09912119060754776
step: 490, loss: 0.161887988448143
step: 500, loss: 0.11230693757534027
step: 510, loss: 0.04886726289987564
step: 520, loss: 0.1113058403134346
step: 530, loss: 0.14934208989143372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9215236346948141, f1=0.9320477502295683,
best_f1=0.9320477502295683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09530947357416153
step: 10, loss: 0.07067200541496277
step: 20, loss: 0.05765102803707123
step: 30, loss: 0.058041878044605255
step: 40, loss: 0.09257400035858154
step: 50, loss: 0.042073290795087814
step: 60, loss: 0.026015039533376694
step: 70, loss: 0.019228406250476837
step: 80, loss: 0.024981185793876648
step: 90, loss: 0.07481984049081802
step: 100, loss: 0.1887667030096054
```

```
step: 110, loss: 0.030407359823584557
step: 120, loss: 0.035178862512111664
step: 130, loss: 0.009071285836398602
step: 140, loss: 0.020195147022604942
step: 150, loss: 0.03386290371417999
step: 160, loss: 0.045531027019023895
step: 170, loss: 0.011994552798569202
step: 180, loss: 0.13284115493297577
step: 190, loss: 0.11682713031768799
step: 200, loss: 0.20328523218631744
step: 210, loss: 0.03636116534471512
step: 220, loss: 0.008271007798612118
step: 230, loss: 0.14648279547691345
step: 240, loss: 0.06075577810406685
step: 250, loss: 0.08856465667486191
step: 260, loss: 0.07997827231884003
step: 270, loss: 0.03145398944616318
step: 280, loss: 0.012170820496976376
step: 290, loss: 0.03869744762778282
step: 300, loss: 0.07970328629016876
step: 310, loss: 0.08881460130214691
step: 320, loss: 0.04585536941885948
step: 330, loss: 0.046421315521001816
step: 340, loss: 0.03632351756095886
step: 350, loss: 0.00241013523191213
step: 360, loss: 0.03609725087881088
step: 370, loss: 0.005334940273314714
step: 380, loss: 0.11968100816011429
step: 390, loss: 0.010620517656207085
step: 400, loss: 0.044932086020708084
step: 410, loss: 0.057147055864334106
step: 420, loss: 0.062923736870288885
step: 430, loss: 0.06203323230147362
step: 440, loss: 0.0113452710211277
step: 450, loss: 0.07722898572683334
step: 460, loss: 0.03785661235451698
step: 470, loss: 0.01319244597107172
step: 480, loss: 0.006824752781540155
step: 490, loss: 0.060314930975437164
step: 500, loss: 0.02341259829699993
step: 510, loss: 0.03620753437280655
step: 520, loss: 0.4005448520183563
step: 530, loss: 0.13425399363040924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9392416628597534, f1=0.9437070938215102,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16528163850307465
step: 10, loss: 0.07783359289169312
step: 20, loss: 0.012463976629078388
step: 30, loss: 0.04424950107932091
step: 40, loss: 0.06761319935321808
step: 50, loss: 0.016772927716374397
step: 60, loss: 0.036038096994161606
step: 70, loss: 0.021110180765390396
step: 80, loss: 0.042676765471696854
step: 90, loss: 0.14416173100471497
step: 100, loss: 0.0450826957821846
step: 110, loss: 0.030227061361074448
step: 120, loss: 0.25710931420326233
step: 130, loss: 0.11204958707094193
step: 140, loss: 0.02994256466627121
step: 150, loss: 0.05501078441739082
step: 160, loss: 0.022705327719449997
step: 170, loss: 0.008044126443564892
step: 180, loss: 0.014018312096595764
step: 190, loss: 0.04097438231110573
step: 200, loss: 0.01341917458921671
step: 210, loss: 0.04040946066379547
step: 220, loss: 0.07683549076318741
step: 230, loss: 0.014244201593101025
step: 240, loss: 0.06864519417285919
step: 250, loss: 0.16876403987407684
step: 260, loss: 0.12194231897592545
step: 270, loss: 0.009518188424408436
step: 280, loss: 0.05749274045228958
step: 290, loss: 0.0036028840113431215
step: 300, loss: 0.2138606756925583
step: 310, loss: 0.0434534065425396
step: 320, loss: 0.038597650825977325
step: 330, loss: 0.015005183406174183
step: 340, loss: 0.007284839171916246
step: 350, loss: 0.053729549050331116
step: 360, loss: 0.009177004918456078
step: 370, loss: 0.03188760578632355
step: 380, loss: 0.012494947761297226
step: 390, loss: 0.02436397597193718
step: 400, loss: 0.062167368829250336
```

```
step: 410, loss: 0.04603840783238411
step: 420, loss: 0.013767914846539497
step: 430, loss: 0.031433433294296265
step: 440, loss: 0.14928029477596283
step: 450, loss: 0.038284655660390854
step: 460, loss: 0.1058514416217804
step: 470, loss: 0.005008346866816282
step: 480, loss: 0.1395040601491928
step: 490, loss: 0.0361703559756279
step: 500, loss: 0.005854157265275717
step: 510, loss: 0.07035180181264877
step: 520, loss: 0.0048277899622917175
step: 530, loss: 0.010282663628458977
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.944649446494465, f1=0.9458583988894032,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004820891655981541
step: 10, loss: 0.002419366268441081
step: 20, loss: 0.04523107409477234
step: 30, loss: 0.10111210495233536
step: 40, loss: 0.007063242141157389
step: 50, loss: 0.09758441895246506
step: 60, loss: 0.0275693815520152092
step: 70, loss: 0.005944686476141214
step: 80, loss: 0.14676491916179657
step: 90, loss: 0.13781827688217163
step: 100, loss: 0.03262699767947197
step: 110, loss: 0.009928819723427296
step: 120, loss: 0.03271494805812836
step: 130, loss: 0.022756710648536682
step: 140, loss: 0.06164459511637688
step: 150, loss: 0.00391466123983264
step: 160, loss: 0.0076707880944013596
step: 170, loss: 0.01427388470619917
step: 180, loss: 0.07076893001794815
step: 190, loss: 0.03770297393202782
step: 200, loss: 0.07505343109369278
step: 210, loss: 0.016590217128396034
step: 220, loss: 0.06890047341585159
```

```
step: 230, loss: 0.018957320600748062
step: 240, loss: 0.001249957480467856
step: 250, loss: 0.17032477259635925
step: 260, loss: 0.002641890896484256
step: 270, loss: 0.029925405979156494
step: 280, loss: 0.005773170385509729
step: 290, loss: 0.026160728186368942
step: 300, loss: 0.007627406157553196
step: 310, loss: 0.007389754056930542
step: 320, loss: 0.12059132009744644
step: 330, loss: 0.006406798958778381
step: 340, loss: 0.0022565166000276804
step: 350, loss: 0.10935691744089127
step: 360, loss: 0.014199437573552132
step: 370, loss: 0.016276506707072258
step: 380, loss: 0.007234628777951002
step: 390, loss: 0.0005502434214577079
step: 400, loss: 0.020638836547732353
step: 410, loss: 0.017706407234072685
step: 420, loss: 0.0260443277657032
step: 430, loss: 0.015622028149664402
step: 440, loss: 0.005647913087159395
step: 450, loss: 0.019009103998541832
step: 460, loss: 0.09455680847167969
step: 470, loss: 0.013141145929694176
step: 480, loss: 0.009657126851379871
step: 490, loss: 0.03225094452500343
step: 500, loss: 0.10082236677408218
step: 510, loss: 0.23668844997882843
step: 520, loss: 0.053155846893787384
step: 530, loss: 0.0788939893245697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9447565543071161, f1=0.9436619718309859,
best_f1=0.9436619718309859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027181566692888737
step: 10, loss: 0.017077812924981117
step: 20, loss: 0.0015306740533560514
step: 30, loss: 0.017692729830741882
step: 40, loss: 0.0007382850162684917
```

```
step: 50, loss: 0.16547709703445435
step: 60, loss: 0.026420513167977333
step: 70, loss: 0.0029300202149897814
step: 80, loss: 0.02916276641190052
step: 90, loss: 0.19400548934936523
step: 100, loss: 0.14892612397670746
step: 110, loss: 0.02498849667608738
step: 120, loss: 0.17957466840744019
step: 130, loss: 0.020064203068614006
step: 140, loss: 0.09840628504753113
step: 150, loss: 0.048633430153131485
step: 160, loss: 0.027115311473608017
step: 170, loss: 0.0785466730594635
step: 180, loss: 0.0161641426384449
step: 190, loss: 0.007945963181555271
step: 200, loss: 0.003538072109222412
step: 210, loss: 0.08449681848287582
step: 220, loss: 0.011843563057482243
step: 230, loss: 0.003129810094833374
step: 240, loss: 0.018710941076278687
step: 250, loss: 0.2694815397262573
step: 260, loss: 0.0037520350888371468
step: 270, loss: 0.01637108437716961
step: 280, loss: 0.06320738792419434
step: 290, loss: 0.010685022920370102
step: 300, loss: 0.025819335132837296
step: 310, loss: 0.04890121892094612
step: 320, loss: 0.005103105213493109
step: 330, loss: 0.05202975496649742
step: 340, loss: 0.007414850872009993
step: 350, loss: 0.000649857334792614
step: 360, loss: 0.004747775383293629
step: 370, loss: 0.0012628078693524003
step: 380, loss: 0.006098945625126362
step: 390, loss: 0.022865012288093567
step: 400, loss: 0.046775221824645996
step: 410, loss: 0.07365431636571884
step: 420, loss: 0.14833050966262817
step: 430, loss: 0.10375937074422836
step: 440, loss: 0.006904118694365025
step: 450, loss: 0.011639057658612728
step: 460, loss: 0.022051963955163956
step: 470, loss: 0.17103822529315948
step: 480, loss: 0.03572603315114975
step: 490, loss: 0.024454470723867416
step: 500, loss: 0.0010859548347070813
step: 510, loss: 0.0043349675834178925
step: 520, loss: 0.10853602737188339
```

```
step: 530, loss: 0.00239013135433197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9464788732394367, f1=0.9443665264142123,
best_f1=0.9443665264142123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03961683437228203
step: 10, loss: 0.003952543716877699
step: 20, loss: 0.02102707140147686
step: 30, loss: 0.001682446338236332
step: 40, loss: 0.033670470118522644
step: 50, loss: 0.0007786196074448526
step: 60, loss: 0.1861191987991333
step: 70, loss: 0.0019729200284928083
step: 80, loss: 0.0029162645805627108
step: 90, loss: 0.008839423768222332
step: 100, loss: 0.0024472225923091173
step: 110, loss: 0.027056632563471794
step: 120, loss: 0.011477435007691383
step: 130, loss: 0.004693569149821997
step: 140, loss: 0.0587516650557518
step: 150, loss: 0.00040771104977466166
step: 160, loss: 0.06887985020875931
step: 170, loss: 0.001091022975742817
step: 180, loss: 0.043818820267915726
step: 190, loss: 0.2664351761341095
step: 200, loss: 0.0034162390511482954
step: 210, loss: 0.0059764632023870945
step: 220, loss: 0.009303324855864048
step: 230, loss: 0.044535838067531586
step: 240, loss: 0.02429545484483242
step: 250, loss: 0.07707346975803375
step: 260, loss: 0.0005607686471194029
step: 270, loss: 0.03135332837700844
step: 280, loss: 0.10007211565971375
step: 290, loss: 0.0022957134060561657
step: 300, loss: 0.006641200743615627
step: 310, loss: 0.06106705591082573
step: 320, loss: 0.0001973812613869086
step: 330, loss: 0.0002645065833348781
step: 340, loss: 0.00032541374093852937
```

```
step: 350, loss: 0.0018565967911854386
step: 360, loss: 0.007650585845112801
step: 370, loss: 0.014884109608829021
step: 380, loss: 0.00262609776109457
step: 390, loss: 0.00022470405383501202
step: 400, loss: 0.0015588854439556599
step: 410, loss: 0.013866853900253773
step: 420, loss: 0.001951699610799551
step: 430, loss: 0.001632632571272552
step: 440, loss: 0.017111124470829964
step: 450, loss: 0.06307432055473328
step: 460, loss: 0.0003829297493211925
step: 470, loss: 0.0009955898858606815
step: 480, loss: 0.0009381602867506444
step: 490, loss: 0.006400598678737879
step: 500, loss: 0.0019780960865318775
step: 510, loss: 0.3055906891822815
step: 520, loss: 0.0012666641268879175
step: 530, loss: 0.014025297947227955
/usr/local/lib/python3.7/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=0.9411214953271028, f1=0.9373831775700935,
best_f1=0.9443665264142123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02614821493625641
step: 10, loss: 0.016586072742938995
step: 20, loss: 0.007249540649354458
step: 30, loss: 0.011922097764909267
step: 40, loss: 0.0018272142624482512
step: 50, loss: 0.01626073755323887
step: 60, loss: 0.00202656676992774
step: 70, loss: 0.0010641837725415826
step: 80, loss: 0.01879049651324749
step: 90, loss: 0.0004905657260678709
step: 100, loss: 0.0020482977852225304
step: 110, loss: 0.0005361589137464762
step: 120, loss: 0.03257128968834877
step: 130, loss: 0.0008446102146990597
step: 140, loss: 0.004641254432499409
step: 150, loss: 0.008807715959846973
step: 160, loss: 0.0003800633130595088
step: 170, loss: 0.0026766464579850435
step: 180, loss: 0.061809394508600235
step: 190, loss: 0.00890466570854187
step: 200, loss: 0.00028188544092699885
```

```
step: 210, loss: 0.0011738564353436232
step: 220, loss: 0.13718543946743011
step: 230, loss: 0.0007650966290384531
step: 240, loss: 0.1456441730260849
step: 250, loss: 0.028507186099886894
step: 260, loss: 0.012983080931007862
step: 270, loss: 0.011015736497938633
step: 280, loss: 0.041560471057891846
step: 290, loss: 0.010139286518096924
step: 300, loss: 0.0009793622884899378
step: 310, loss: 0.006360391154885292
step: 320, loss: 0.028878364711999893
step: 330, loss: 0.0004235812521073967
step: 340, loss: 0.0006079430459067225
step: 350, loss: 0.006723240949213505
step: 360, loss: 0.014678302221000195
step: 370, loss: 0.03844686970114708
step: 380, loss: 0.003884692909196019
step: 390, loss: 0.005771899130195379
step: 400, loss: 0.008710035122931004
step: 410, loss: 0.00017598156409803778
step: 420, loss: 0.0028896748553961515
step: 430, loss: 0.013960241340100765
step: 440, loss: 0.000319989281706512
step: 450, loss: 0.00957751739770174
step: 460, loss: 0.004566890187561512
step: 470, loss: 0.1335035115480423
step: 480, loss: 0.0025562888476997614
step: 490, loss: 0.03416156768798828
step: 500, loss: 0.07362130284309387
step: 510, loss: 0.0038005104288458824
step: 520, loss: 0.0007973626488819718
step: 530, loss: 0.0058555277064442635
/usr/local/lib/python3.7/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=0.9442100328176277, f1=0.9435897435897437,
best_f1=0.9443665264142123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036419571842998266
step: 10, loss: 0.03560146316885948
step: 20, loss: 0.000287647417280823
step: 30, loss: 0.00028783874586224556
step: 40, loss: 9.217159094987437e-05
step: 50, loss: 0.0001274755923077464
step: 60, loss: 0.00021957876742817461
```

```
step: 70, loss: 0.001407222356647253
step: 80, loss: 0.029296040534973145
step: 90, loss: 0.00016026697994675487
step: 100, loss: 0.0002676607109606266
step: 110, loss: 0.00020648531790357083
step: 120, loss: 0.00038429233245551586
step: 130, loss: 9.724726260174066e-05
step: 140, loss: 0.023384710773825645
step: 150, loss: 0.0005668115918524563
step: 160, loss: 0.002735392889007926
step: 170, loss: 0.02904951199889183
step: 180, loss: 0.056545209139585495
step: 190, loss: 0.00876458827406168
step: 200, loss: 0.004366111010313034
step: 210, loss: 0.0693199634552002
step: 220, loss: 0.0003037629066966474
step: 230, loss: 0.17742492258548737
step: 240, loss: 0.02000650018453598
step: 250, loss: 0.002492735628038645
step: 260, loss: 0.00026958322268910706
step: 270, loss: 0.01659495197236538
step: 280, loss: 0.007063181605190039
step: 290, loss: 0.004350604489445686
step: 300, loss: 0.0003805679443757981
step: 310, loss: 0.0024415848311036825
step: 320, loss: 0.00037927820812910795
step: 330, loss: 0.0008364531677216291
step: 340, loss: 0.059173453599214554
step: 350, loss: 0.00012032323866151273
step: 360, loss: 0.16430389881134033
step: 370, loss: 0.056507330387830734
step: 380, loss: 0.0015477873384952545
step: 390, loss: 0.004625044763088226
step: 400, loss: 0.008932794444262981
step: 410, loss: 0.000936671975068748
step: 420, loss: 6.779766408726573e-05
step: 430, loss: 0.027951300144195557
step: 440, loss: 0.016424529254436493
step: 450, loss: 0.0006168322288431227
step: 460, loss: 0.0018647827673703432
step: 470, loss: 0.052200306206941605
step: 480, loss: 0.11472873389720917
step: 490, loss: 0.002870651427656412
step: 500, loss: 0.0018399322871118784
step: 510, loss: 0.35790783166885376
step: 520, loss: 0.00010020067566074431
step: 530, loss: 0.01226902473717928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9465791940018744, f1=0.9400187441424556,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002882523462176323
step: 10, loss: 0.010894796811044216
step: 20, loss: 6.777418457204476e-05
step: 30, loss: 0.05759013816714287
step: 40, loss: 0.00013028366083744913
step: 50, loss: 0.0094906622543931
step: 60, loss: 0.00012665687245316803
step: 70, loss: 0.003988600801676512
step: 80, loss: 0.0037871815729886293
step: 90, loss: 0.20359419286251068
step: 100, loss: 0.009880981408059597
step: 110, loss: 0.008794025518000126
step: 120, loss: 0.0010652265045791864
step: 130, loss: 0.00024543129256926477
step: 140, loss: 0.022234836593270302
step: 150, loss: 0.003758043982088566
step: 160, loss: 0.006655079312622547
step: 170, loss: 0.0088646300137043
step: 180, loss: 0.004199989605695009
step: 190, loss: 0.06675729155540466
step: 200, loss: 0.008364824578166008
step: 210, loss: 0.007699937559664249
step: 220, loss: 0.000548575830180198
step: 230, loss: 0.009410443715751171
step: 240, loss: 0.0007671342464163899
step: 250, loss: 0.0028632478788495064
step: 260, loss: 0.0023168842308223248
step: 270, loss: 0.002743504708632827
step: 280, loss: 0.010910524986684322
step: 290, loss: 0.00019409281958360225
step: 300, loss: 0.03570529446005821
step: 310, loss: 0.0005218322621658444
step: 320, loss: 0.00041693111415952444
step: 330, loss: 0.001232035574503243
step: 340, loss: 0.002193338703364134
step: 350, loss: 0.014994683675467968
step: 360, loss: 0.0019235991640016437
```

```
step: 370, loss: 0.0008540611597709358
step: 380, loss: 0.0008849096484482288
step: 390, loss: 0.00015482692106161267
step: 400, loss: 0.0017417182680219412
step: 410, loss: 0.021062131971120834
step: 420, loss: 0.003525547683238983
step: 430, loss: 0.05264924839138985
step: 440, loss: 0.00016597214562352747
step: 450, loss: 0.005005622748285532
step: 460, loss: 0.0009619450429454446
step: 470, loss: 0.0008715060539543629
step: 480, loss: 0.0013511341530829668
step: 490, loss: 0.0019905241206288338
step: 500, loss: 0.0007388555677607656
step: 510, loss: 0.05311049520969391
step: 520, loss: 0.004102956969290972
step: 530, loss: 0.008308112621307373
/usr/local/lib/python3.7/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=0.9316163410301954, f1=0.9385925593904079,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03199721872806549
step: 10, loss: 0.002629654947668314
step: 20, loss: 0.0002096926182275638
step: 30, loss: 0.010177145712077618
step: 40, loss: 0.012025970965623856
step: 50, loss: 0.006891579832881689
step: 60, loss: 0.0109678003937006
step: 70, loss: 0.00035845045931637287
step: 80, loss: 0.003227213630452752
step: 90, loss: 0.002408646047115326
step: 100, loss: 0.0024387489538639784
step: 110, loss: 0.01606808416545391
step: 120, loss: 0.0004541784292086959
step: 130, loss: 0.1279468834400177
step: 140, loss: 0.00024180064792744815
step: 150, loss: 0.0021979461889714003
step: 160, loss: 0.017267446964979172
step: 170, loss: 0.04528653249144554
step: 180, loss: 0.00021772657055407763
step: 190, loss: 5.337666152627207e-05
step: 200, loss: 0.00014690705575048923
step: 210, loss: 0.016013098880648613
step: 220, loss: 0.00126259948592633
```

```
step: 230, loss: 0.0008926928276196122
step: 240, loss: 8.653740951558575e-05
step: 250, loss: 0.017907002940773964
step: 260, loss: 0.23706060647964478
step: 270, loss: 0.0013445188524201512
step: 280, loss: 0.034908976405858994
step: 290, loss: 0.00031251518521145791
step: 300, loss: 0.002127181040123105
step: 310, loss: 0.026772387325763702
step: 320, loss: 0.06462706625461578
step: 330, loss: 0.0020139191765338182
step: 340, loss: 0.0005212083924561739
step: 350, loss: 0.0002510453632567078
step: 360, loss: 0.0001192955969600007
step: 370, loss: 0.0031971668358892202
step: 380, loss: 0.0017904005944728851
step: 390, loss: 0.00034375692484900355
step: 400, loss: 8.996998076327145e-05
step: 410, loss: 0.0026117803063243628
step: 420, loss: 7.148356235120445e-05
step: 430, loss: 0.00011833674943773076
step: 440, loss: 0.003463735803961754
step: 450, loss: 0.011261086910963058
step: 460, loss: 0.000409464817494154
step: 470, loss: 0.01132083497941494
step: 480, loss: 0.0011934692738577724
step: 490, loss: 0.0006587860989384353
step: 500, loss: 0.00017968406609725207
step: 510, loss: 7.88888064562343e-05
step: 520, loss: 0.0021574972197413445
step: 530, loss: 0.03185446187853813
/usr/local/lib/python3.7/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=0.9396396396396396, f1=0.9392165691130121,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10114318877458572
step: 10, loss: 0.0008215405396185815
step: 20, loss: 0.003205780638381839
step: 30, loss: 0.00048582087038084865
step: 40, loss: 0.0020692613907158375
step: 50, loss: 0.0016915975138545036
step: 60, loss: 0.0027128765359520912
step: 70, loss: 0.0008364471141248941
step: 80, loss: 0.0030457302927970886
```

```
step: 90, loss: 0.00016688188770785928
step: 100, loss: 0.01144933607429266
step: 110, loss: 0.007070411927998066
step: 120, loss: 0.001981711946427822
step: 130, loss: 0.00024395259970333427
step: 140, loss: 0.0005195661797188222
step: 150, loss: 0.0014332669088616967
step: 160, loss: 0.00015957992582116276
step: 170, loss: 0.0041606673039495945
step: 180, loss: 0.00021843871218152344
step: 190, loss: 0.0011201187735423446
step: 200, loss: 0.00027464295271784067
step: 210, loss: 5.4894400818739086e-05
step: 220, loss: 0.06786130368709564
step: 230, loss: 0.14049994945526123
step: 240, loss: 0.01203138381242752
step: 250, loss: 0.0012868390185758471
step: 260, loss: 0.0005325779202394187
step: 270, loss: 0.010539481416344643
step: 280, loss: 0.001433982397429645
step: 290, loss: 0.00014142545114737004
step: 300, loss: 0.00018343109695706517
step: 310, loss: 0.09192254394292831
step: 320, loss: 0.0011025051353499293
step: 330, loss: 0.00011008380533894524
step: 340, loss: 0.01356184110045433
step: 350, loss: 0.0001085176772903651
step: 360, loss: 0.0016812359681352973
step: 370, loss: 0.0002390518202446401
step: 380, loss: 0.00038341316394507885
step: 390, loss: 0.0003407049516681582
step: 400, loss: 0.0012357741361483932
step: 410, loss: 0.004204131662845612
step: 420, loss: 0.0017996453680098057
step: 430, loss: 0.0011541671119630337
step: 440, loss: 0.00017614651005715132
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 450, loss: 0.000651300884783268
step: 460, loss: 0.00023394175514113158
step: 470, loss: 5.715718725696206e-05
step: 480, loss: 0.0008123706211335957
step: 490, loss: 0.004132261965423822
step: 500, loss: 0.036585260182619095
step: 510, loss: 0.0022259464021772146
step: 520, loss: 9.279026562580839e-05
step: 530, loss: 0.0024734761100262403
/usr/local/lib/python3.7/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=0.9444184960298926, f1=0.9446254071661238,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011019419180229306
step: 10, loss: 0.010433804243803024
step: 20, loss: 0.0012653071898967028
step: 30, loss: 5.226701250649057e-05
step: 40, loss: 0.0006200637435540557
step: 50, loss: 0.0015078127617016435
step: 60, loss: 0.0032513842452317476
step: 70, loss: 0.0005847160355187953
step: 80, loss: 3.72380563931074e-05
step: 90, loss: 0.0004061970394104719
step: 100, loss: 0.017430752515792847
step: 110, loss: 0.0006814278312958777
step: 120, loss: 5.3645941079594195e-05
step: 130, loss: 0.004791536834090948
step: 140, loss: 0.00031296417000703514
step: 150, loss: 4.329564035288058e-05
step: 160, loss: 0.01126316748559475
step: 170, loss: 0.000478212081361562
step: 180, loss: 1.9680413970490918e-05
step: 190, loss: 0.0002560480497777462
step: 200, loss: 0.00042913234210573137
step: 210, loss: 2.908932037826162e-05
step: 220, loss: 0.024870403110980988
step: 230, loss: 0.0021953831892460585
step: 240, loss: 0.0008991178474389017
step: 250, loss: 3.299726085970178e-05
step: 260, loss: 0.0006822231807745993
step: 270, loss: 0.14049378037452698
step: 280, loss: 8.016404899535701e-05
step: 290, loss: 0.0001358169101877138
step: 300, loss: 0.00723667535930872
step: 310, loss: 0.006536294240504503
step: 320, loss: 0.0012128339149057865
step: 330, loss: 0.1914299726486206
step: 340, loss: 0.001391438185237348
step: 350, loss: 0.00022297466057352722
step: 360, loss: 0.010951253585517406
step: 370, loss: 0.000979563919827342
step: 380, loss: 0.003422216512262821
step: 390, loss: 0.08256827294826508
step: 400, loss: 0.00035702227614820004
step: 410, loss: 0.02917839214205742
```

```
step: 420, loss: 0.062057413160800934
step: 430, loss: 0.0038449319545179605
step: 440, loss: 0.01728726550936699
step: 450, loss: 0.005751064047217369
step: 460, loss: 0.0006409060442820191
step: 470, loss: 0.0049041821621358395
step: 480, loss: 0.0016734832897782326
step: 490, loss: 0.0017670539673417807
step: 500, loss: 0.00047899025958031416
step: 510, loss: 0.0043299198150634766
step: 520, loss: 0.00020895313355140388
step: 530, loss: 0.0006914610275998712
/usr/local/lib/python3.7/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=0.9461862423958821, f1=0.9421101774042949,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.589590728050098e-05
step: 10, loss: 0.00204155477769516706
step: 20, loss: 0.00012276737834326923
step: 30, loss: 0.0005410315352492034
step: 40, loss: 0.0023865592665970325
step: 50, loss: 0.005601843353360891
step: 60, loss: 0.0022918188478797674
step: 70, loss: 0.001295703463256359
step: 80, loss: 0.0005319819319993258
step: 90, loss: 9.074641275219619e-05
step: 100, loss: 0.003007502993568778
step: 110, loss: 0.00014675466809421778
step: 120, loss: 0.0005504540167748928
step: 130, loss: 0.0007113323081284761
step: 140, loss: 0.0001606596342753619
step: 150, loss: 0.0002919812686741352
step: 160, loss: 0.009435958229005337
step: 170, loss: 0.05248808115720749
step: 180, loss: 0.000836468767374754
step: 190, loss: 0.00024334611953236163
step: 200, loss: 9.174096339847893e-05
step: 210, loss: 0.0001414678990840912
step: 220, loss: 0.0002594218240119517
step: 230, loss: 0.0020870263688266277
step: 240, loss: 8.643612818559632e-05
step: 250, loss: 0.0002352725132368505
step: 260, loss: 0.001653697807341814
step: 270, loss: 0.0006713686743751168
```

```
step: 280, loss: 0.0007354682311415672
step: 290, loss: 0.1507011502981186
step: 300, loss: 0.0011423235991969705
step: 310, loss: 0.00011096634261775762
step: 320, loss: 0.0008942185086198151
step: 330, loss: 0.00144218560308218
step: 340, loss: 0.001343936426565051
step: 350, loss: 4.6970119001343846e-05
step: 360, loss: 0.05748412758111954
step: 370, loss: 0.0018387231975793839
step: 380, loss: 0.0003415377577766776
step: 390, loss: 9.597774624126032e-05
step: 400, loss: 0.00045352106099016964
step: 410, loss: 9.8856893600896e-05
step: 420, loss: 0.0037474280688911676
step: 430, loss: 0.00014449184527620673
step: 440, loss: 6.134670547908172e-05
step: 450, loss: 0.02169562131166458
step: 460, loss: 0.007602231577038765
step: 470, loss: 0.0003433080273680389
step: 480, loss: 2.95554655167507e-05
step: 490, loss: 0.0002834201150108129
step: 500, loss: 0.00029773544520139694
step: 510, loss: 0.0001678693079156801
step: 520, loss: 0.00015937983698677272
step: 530, loss: 3.717965228133835e-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=0.9464368886818817, f1=0.9451672862453532,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017352236900478601
step: 10, loss: 0.000651594833470881
step: 20, loss: 0.0010849593672901392
step: 30, loss: 0.010357976891100407
step: 40, loss: 0.0001311331579927355
step: 50, loss: 0.00569720147177577
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 60, loss: 0.0003215954639017582
step: 70, loss: 9.631737339077517e-05
step: 80, loss: 0.005442284047603607
step: 90, loss: 1.5076040654093958e-05
step: 100, loss: 0.001189833739772439
step: 110, loss: 0.00027754725306294860
step: 120, loss: 2.4366414436371997e-05
```

```
step: 130, loss: 0.00010310401557944715
step: 140, loss: 0.0014694822020828724
step: 150, loss: 9.452831727685407e-05
step: 160, loss: 0.002266465686261654
step: 170, loss: 8.913165947888047e-05
step: 180, loss: 0.00014793888840358704
step: 190, loss: 0.010656560771167278
step: 200, loss: 0.0003643041127361357
step: 210, loss: 0.0005499255494214594
step: 220, loss: 3.4137690818170086e-05
step: 230, loss: 0.0011016640346497297
step: 240, loss: 0.00046230602310970426
step: 250, loss: 0.0009866309119388461
step: 260, loss: 0.0017361062346026301
step: 270, loss: 0.0004122614045627415
step: 280, loss: 0.0003505891945678741
step: 290, loss: 6.618897168664262e-05
step: 300, loss: 0.002836809027940035
step: 310, loss: 0.005742188543081284
step: 320, loss: 2.49848762905458e-05
step: 330, loss: 7.729460048722103e-05
step: 340, loss: 9.721230890136212e-05
step: 350, loss: 0.0002677152515389025
step: 360, loss: 0.000560897751711309
step: 370, loss: 0.00039133161772042513
step: 380, loss: 0.04769738018512726
step: 390, loss: 0.0017688765656203032
step: 400, loss: 0.010580958798527718
step: 410, loss: 0.00010954867320833728
step: 420, loss: 0.0012635914608836174
step: 430, loss: 0.001115859835408628
step: 440, loss: 0.01834775321185589
step: 450, loss: 4.194331995677203e-05
step: 460, loss: 0.179985910654068
step: 470, loss: 2.6292125767213292e-05
step: 480, loss: 0.0011457569198682904
step: 490, loss: 0.00011528508912306279
step: 500, loss: 0.0013550681760534644
step: 510, loss: 0.023491468280553818
step: 520, loss: 8.110314229270443e-05
step: 530, loss: 6.63043319946155e-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=0.9422018348623853, f1=0.9410150891632373,
best_f1=0.9400187441424556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002702204801607877
step: 10, loss: 0.00020533375209197402
step: 20, loss: 0.004521501716226339
step: 30, loss: 0.009911488741636276
step: 40, loss: 0.00030710865394212306
step: 50, loss: 0.0011353912996128201
step: 60, loss: 0.010301565751433372
step: 70, loss: 9.926026541506872e-05
step: 80, loss: 0.0003105357172898948
step: 90, loss: 0.0002408101863693446
step: 100, loss: 8.161198638845235e-05
step: 110, loss: 0.017956266179680824
step: 120, loss: 0.0005048227030783892
step: 130, loss: 0.0016928408294916153
step: 140, loss: 0.00031521497294306755
step: 150, loss: 0.0005039394600316882
step: 160, loss: 0.00012880384747404605
step: 170, loss: 6.037829734850675e-05
step: 180, loss: 0.009696139954030514
step: 190, loss: 0.0014868525322526693
step: 200, loss: 0.000844186230096966
step: 210, loss: 0.0006357667734846473
step: 220, loss: 0.002336242701858282
step: 230, loss: 0.0001035758396028541
step: 240, loss: 0.00011437795910751447
step: 250, loss: 9.286741260439157e-05
step: 260, loss: 0.0004555558552965522
step: 270, loss: 4.3571784772211686e-05
step: 280, loss: 6.551775004481897e-05
step: 290, loss: 5.2612107538152486e-05
step: 300, loss: 6.237822526600212e-05
step: 310, loss: 0.0005871852044947445
step: 320, loss: 0.0006164571968838573
step: 330, loss: 4.5273023715708405e-05
step: 340, loss: 0.00044622557470574975
step: 350, loss: 0.00694569107145071
step: 360, loss: 0.000313961150823161
step: 370, loss: 0.0011199311120435596
step: 380, loss: 0.00011942745913984254
step: 390, loss: 0.00028676603687927127
step: 400, loss: 0.04087729752063751
step: 410, loss: 0.0001615914807189256
step: 420, loss: 7.42885094950907e-05
step: 430, loss: 7.447822281392291e-05
step: 440, loss: 0.0003801952989306301
step: 450, loss: 0.001242531230673194
step: 460, loss: 0.006691050250083208
```

```
step: 470, loss: 0.0005478246021084487
step: 480, loss: 0.004022802226245403
step: 490, loss: 0.00024361982650589198
step: 500, loss: 0.0024744688998907804
step: 510, loss: 4.467116013984196e-05
step: 520, loss: 9.467341442359611e-05
step: 530, loss: 0.0012554212007671595
/usr/local/lib/python3.7/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=0.9430588235294118, f1=0.9433255269320844,
best_f1=0.9400187441424556
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:22, 252.70it/s]
load_f1 = 0.9480093676814988
real_f1 = 0.9460853258321612
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, 196.44it/s]

4.1.4 Amazon-Google - Running the matcher

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


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

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

Selected optimization level 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

400

```
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: 0.5023806691169739
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.4619891047477722
step: 20, loss: 0.45627763867378235
step: 30, loss: 0.3132309913635254
step: 40, loss: 0.30115076899528503
step: 50, loss: 0.47457370162010193
step: 60, loss: 0.5139436721801758
step: 70, loss: 0.3154652416706085
step: 80, loss: 0.35674452781677246
step: 90, loss: 0.24780718982219696
step: 100, loss: 0.2391887605190277
step: 110, loss: 0.2919805943965912
step: 120, loss: 0.4304504990577698
step: 130, loss: 0.2557395398616791
step: 140, loss: 0.4535140097141266
step: 150, loss: 0.359980046749115
step: 160, loss: 0.5069124102592468
step: 170, loss: 0.24713388085365295

```
step: 180, loss: 0.39014166593551636
step: 190, loss: 0.6391063928604126
step: 200, loss: 0.3691266179084778
step: 210, loss: 0.4980955719947815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3970491290092468
step: 10, loss: 0.1765946000814438
step: 20, loss: 0.48139968514442444
step: 30, loss: 0.49913039803504944
step: 40, loss: 0.42836299538612366
step: 50, loss: 0.3929409086704254
step: 60, loss: 0.32894501090049744
step: 70, loss: 0.315377414226532
step: 80, loss: 0.292903870344162
step: 90, loss: 0.29144713282585144
step: 100, loss: 0.426622211933136
step: 110, loss: 0.3023383319377899
step: 120, loss: 0.20361903309822083
step: 130, loss: 0.1885954737663269
step: 140, loss: 0.17474989593029022
step: 150, loss: 0.4377463459968567
step: 160, loss: 0.13201454281806946
step: 170, loss: 0.6467089056968689
step: 180, loss: 0.29974091053009033
step: 190, loss: 0.3147304058074951
step: 200, loss: 0.1364554613828659
step: 210, loss: 0.295105516910553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5129151291512914, f1=0.5289855072463767,
best_f1=0.5289855072463767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1888449788093567
step: 10, loss: 0.12082573771476746
step: 20, loss: 0.40158283710479736
step: 30, loss: 0.12853704392910004
step: 40, loss: 0.29633423686027527
step: 50, loss: 0.27721649408340454
step: 60, loss: 0.2622973322868347
step: 70, loss: 0.1754002422094345
step: 80, loss: 0.3015010952949524
step: 90, loss: 0.19999182224273682
step: 100, loss: 0.2538784146308899
step: 110, loss: 0.10652996599674225
step: 120, loss: 0.12046628445386887
step: 130, loss: 0.34014835953712463
step: 140, loss: 0.2726033329963684
step: 150, loss: 0.2601395547389984
step: 160, loss: 0.10198428481817245
step: 170, loss: 0.22896593809127808
step: 180, loss: 0.05844448506832123
step: 190, loss: 0.06185244396328926
step: 200, loss: 0.1258523017168045
step: 210, loss: 0.21319101750850677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5759368836291914, f1=0.570264765784114,
best_f1=0.570264765784114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10241886973381042
step: 10, loss: 0.17212866246700287
step: 20, loss: 0.08972187340259552
step: 30, loss: 0.14222247898578644
step: 40, loss: 0.05728193372488022
step: 50, loss: 0.11216951906681061
step: 60, loss: 0.1893225461244583
step: 70, loss: 0.11213500797748566
step: 80, loss: 0.17784593999385834
step: 90, loss: 0.15480995178222656
step: 100, loss: 0.235197976231575
step: 110, loss: 0.3324156701564789

```
step: 120, loss: 0.29226192831993103
step: 130, loss: 0.12596739828586578
step: 140, loss: 0.5807802081108093
step: 150, loss: 0.13381628692150116
step: 160, loss: 0.13466066122055054
step: 170, loss: 0.11180194467306137
step: 180, loss: 0.04692547395825386
step: 190, loss: 0.18540608882904053
step: 200, loss: 0.1247270330786705
step: 210, loss: 0.2962329387664795
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.5454545454545454,
best_f1=0.570264765784114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17383477091789246
step: 10, loss: 0.11274617910385132
step: 20, loss: 0.1741020381450653
step: 30, loss: 0.09795920550823212
step: 40, loss: 0.22913825511932373
step: 50, loss: 0.185018852353096
step: 60, loss: 0.1881033331155777
step: 70, loss: 0.12445273995399475
step: 80, loss: 0.1446131318807602
step: 90, loss: 0.16279791295528412
step: 100, loss: 0.016281209886074066
step: 110, loss: 0.05472395569086075
step: 120, loss: 0.15153497457504272
step: 130, loss: 0.16513963043689728
step: 140, loss: 0.15137845277786255
step: 150, loss: 0.08570334315299988
step: 160, loss: 0.07302475720643997
step: 170, loss: 0.046110570430755615
step: 180, loss: 0.18455290794372559
step: 190, loss: 0.17622323334217072
step: 200, loss: 0.228658527135849
step: 210, loss: 0.13418090343475342
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.600375234521576, f1=0.5826771653543307,
```

```
best_f1=0.5826771653543307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08624430745840073
step: 10, loss: 0.08821871876716614
step: 20, loss: 0.05570363998413086
step: 30, loss: 0.06534942239522934
step: 40, loss: 0.1134813129901886
step: 50, loss: 0.13767710328102112
step: 60, loss: 0.17612680792808533
step: 70, loss: 0.12020275741815567
step: 80, loss: 0.2396731674671173
step: 90, loss: 0.11211206763982773
step: 100, loss: 0.19879309833049774
step: 110, loss: 0.19697101414203644
step: 120, loss: 0.026860155165195465
step: 130, loss: 0.032481417059898376
step: 140, loss: 0.15586906671524048
step: 150, loss: 0.06242777407169342
step: 160, loss: 0.07440926879644394
step: 170, loss: 0.08627462387084961
step: 180, loss: 0.1619279980659485
step: 190, loss: 0.13188804686069489
step: 200, loss: 0.12846101820468903
step: 210, loss: 0.10241536796092987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6061705989110708, 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: 0.07716929912567139
step: 10, loss: 0.013312255963683128
step: 20, loss: 0.2162865847349167
step: 30, loss: 0.173591315746307737
step: 40, loss: 0.06987406313419342
step: 50, loss: 0.08929473161697388
step: 60, loss: 0.2664963901042938
step: 70, loss: 0.11716267466545105
step: 80, loss: 0.06837276369333267
step: 90, loss: 0.042993783950805664
```

```
step: 100, loss: 0.11610331386327744
step: 110, loss: 0.22648437321186066
step: 120, loss: 0.19530636072158813
step: 130, loss: 0.08014465868473053
step: 140, loss: 0.07833859324455261
step: 150, loss: 0.07661429047584534
step: 160, loss: 0.23705421388149261
step: 170, loss: 0.388604998588562
step: 180, loss: 0.05952734500169754
step: 190, loss: 0.12286382913589478
step: 200, loss: 0.151619091629982
step: 210, loss: 0.11701810359954834
/usr/local/lib/python3.7/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=0.6031746031746031, f1=0.5652173913043479,
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: 0.0558590292930603
step: 10, loss: 0.06059752032160759
step: 20, loss: 0.0601971410214901
step: 30, loss: 0.10099655389785767
step: 40, loss: 0.0434783473610878
step: 50, loss: 0.01779823563992977
step: 60, loss: 0.0226856991648674
step: 70, loss: 0.032139163464307785
step: 80, loss: 0.07402315735816956
step: 90, loss: 0.13289578258991241
step: 100, loss: 0.18331162631511688
step: 110, loss: 0.13188906013965607
step: 120, loss: 0.1342703402042389
step: 130, loss: 0.03324006497859955
step: 140, loss: 0.03343519568443298
step: 150, loss: 0.21493513882160187
step: 160, loss: 0.2535378336906433
step: 170, loss: 0.08906091749668121
step: 180, loss: 0.06336434930562973
step: 190, loss: 0.07407668232917786
step: 200, loss: 0.08183276653289795
step: 210, loss: 0.11858096718788147
/usr/local/lib/python3.7/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=0.5969498910675382, f1=0.581081081081081,
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: 0.09190427511930466
step: 10, loss: 0.10748139768838882
step: 20, loss: 0.13459855318069458
step: 30, loss: 0.0035043740645051003
step: 40, loss: 0.028154492378234863
step: 50, loss: 0.025521699339151382
step: 60, loss: 0.1370960772037506
step: 70, loss: 0.023368002846837044
step: 80, loss: 0.12858329713344574
step: 90, loss: 0.016516732051968575
step: 100, loss: 0.06825205683708191
step: 110, loss: 0.10625215619802475
step: 120, loss: 0.1978062093257904
step: 130, loss: 0.03348090499639511
step: 140, loss: 0.2551603317260742
step: 150, loss: 0.0063626752234995365
step: 160, loss: 0.1616116464138031
step: 170, loss: 0.014512705616652966
step: 180, loss: 0.057753417640924454
step: 190, loss: 0.013645593076944351
step: 200, loss: 0.012645983137190342
step: 210, loss: 0.1451951414346695
/usr/local/lib/python3.7/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=0.5954465849387041, f1=0.5985915492957746,
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: 0.06158706545829773
step: 10, loss: 0.01902494765818119
step: 20, loss: 0.029067106544971466
step: 30, loss: 0.002302213804796338
step: 40, loss: 0.04204665496945381
step: 50, loss: 0.06835037469863892
step: 60, loss: 0.09110257029533386
step: 70, loss: 0.060343995690345764
step: 80, loss: 0.03507538139820099
step: 90, loss: 0.1403123289346695
step: 100, loss: 0.028938043862581253
step: 110, loss: 0.009984295815229416
step: 120, loss: 0.058400750160217285
step: 130, loss: 0.03609021380543709
step: 140, loss: 0.02461232990026474
step: 150, loss: 0.061216361820697784

```
step: 160, loss: 0.03617635369300842
step: 170, loss: 0.09823345392942429
step: 180, loss: 0.0061669000424444675
step: 190, loss: 0.036218978464603424
step: 200, loss: 0.06707427650690079
step: 210, loss: 0.04180959239602089
```
/usr/local/lib/python3.7/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=0.5946969696969697, f1=0.589430894308943,
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: 0.05157003924250603
step: 10, loss: 0.1069047823548317
step: 20, loss: 0.013243775814771652
step: 30, loss: 0.1382347047328949
step: 40, loss: 0.014718125574290752
step: 50, loss: 0.13149338960647583
step: 60, loss: 0.008499585092067719
step: 70, loss: 0.02037058211863041
step: 80, loss: 0.027421072125434875
step: 90, loss: 0.20227886736392975
step: 100, loss: 0.02177973836660385
step: 110, loss: 0.009457387961447239
step: 120, loss: 0.03616940230131149
step: 130, loss: 0.003122388618066907
step: 140, loss: 0.2798064351081848
step: 150, loss: 0.05251626297831535
step: 160, loss: 0.006713980808854103
step: 170, loss: 0.12205825746059418
step: 180, loss: 0.013407967053353786
step: 190, loss: 0.10256771743297577
step: 200, loss: 0.029707347974181175
step: 210, loss: 0.05329803749918938
```
/usr/local/lib/python3.7/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=0.5876288659793815, f1=0.5913043478260869,
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: 0.028520070016384125
step: 10, loss: 0.03647715598344803
step: 20, loss: 0.05524195358157158
step: 30, loss: 0.0012704709079116583
```

```
step: 40, loss: 0.0010724077001214027
step: 50, loss: 0.002851763740181923
step: 60, loss: 0.03938612341880798
step: 70, loss: 0.005393676925450563
step: 80, loss: 0.16237568855285645
step: 90, loss: 0.07304999977350235
step: 100, loss: 0.003533021081238985
step: 110, loss: 0.22974057495594025
step: 120, loss: 0.0028463148046284914
step: 130, loss: 0.005835448857396841
step: 140, loss: 0.06860307604074478
step: 150, loss: 0.003010872285813093
step: 160, loss: 0.029903167858719826
step: 170, loss: 0.033112455159425735
step: 180, loss: 0.016478335484862328
step: 190, loss: 0.018581295385956764
step: 200, loss: 0.00957843940705061
step: 210, loss: 0.04337459057569504
/usr/local/lib/python3.7/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=0.579476861167002, f1=0.5958333333333334,
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: 0.016761621460318565
step: 10, loss: 0.00903010368347168
step: 20, loss: 0.003349423175677657
step: 30, loss: 0.004541751462966204
step: 40, loss: 0.01255915965884924
step: 50, loss: 0.12172802537679672
step: 60, loss: 0.01737326942384243
step: 70, loss: 0.052220650017261505
step: 80, loss: 0.010386127978563309
step: 90, loss: 0.004986660089343786
step: 100, loss: 0.0022915659938007593
step: 110, loss: 0.0589335672557354
step: 120, loss: 0.004548253025859594
step: 130, loss: 0.002806044416502118
step: 140, loss: 0.0024816258810460567
step: 150, loss: 0.00417298823595047
step: 160, loss: 0.008718001656234264
step: 170, loss: 0.09001316130161285
step: 180, loss: 0.006776988040655851
step: 190, loss: 0.020883038640022278
step: 200, loss: 0.17357930541038513
step: 210, loss: 0.012080504558980465
```

```
/usr/local/lib/python3.7/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=0.5980952380952381, f1=0.6090373280943024,
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: 0.012290923822671175
step: 10, loss: 0.05147329717874527
step: 20, loss: 0.009428363293409348
step: 30, loss: 0.10707120597362518
step: 40, loss: 0.0025100749917328358
step: 50, loss: 0.02647021971642971
step: 60, loss: 0.03943072631955147
step: 70, loss: 0.000663531303871423
step: 80, loss: 0.019929783418774605
step: 90, loss: 0.0429009273648262
step: 100, loss: 0.013154236599802971
step: 110, loss: 0.0013478390173986554
step: 120, loss: 0.014405698515474796
step: 130, loss: 0.07081378996372223
step: 140, loss: 0.0021536473650485277
step: 150, loss: 0.05056222528219223
step: 160, loss: 0.033398669213056564
step: 170, loss: 0.017461517825722694
step: 180, loss: 0.0004699309647548944
step: 190, loss: 0.14607860147953033
step: 200, loss: 0.038497138768434525
step: 210, loss: 0.002112105954438448
/usr/local/lib/python3.7/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=0.5708154506437768, f1=0.5964912280701754,
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: 0.002569011878222227
step: 10, loss: 0.0018998612649738789
step: 20, loss: 0.019318385049700737
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.005516088102012873
step: 40, loss: 0.005903883837163448
step: 50, loss: 0.013494358398020267
step: 60, loss: 0.005989151541143656
step: 70, loss: 0.01799202337861061
step: 80, loss: 0.03339271992444992
```

```
step: 90, loss: 0.017477190122008324
step: 100, loss: 0.0034707796294242144
step: 110, loss: 0.010774660855531693
step: 120, loss: 0.0038837764877825975
step: 130, loss: 0.016522562131285667
step: 140, loss: 0.011503576301038265
step: 150, loss: 0.15591397881507874
step: 160, loss: 0.12810669839382172
step: 170, loss: 0.017751168459653854
step: 180, loss: 0.020893186330795288
step: 190, loss: 0.09237697720527649
step: 200, loss: 0.004367291461676359
step: 210, loss: 0.04650839418172836
/usr/local/lib/python3.7/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=0.5758157389635318, f1=0.6099009900990099,
best_f1=0.5714285714285714
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, 488.43it/s]
load_f1 = 0.6122448979591838
real_f1 = 0.6153846153846153
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, 195.31it/s]
```

4.1.5   Walmart-Amazon - Running the matcher

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


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

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

```
Downloading: 100% 481/481 [00:00<00:00, 427kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 795kB/s]
```

```
Downloading: 100% 456k/456k [00:00<00:00, 516kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 69.3MB/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: 0.44697827100753784
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4618367850780487
step: 20, loss: 0.2513604462146759
step: 30, loss: 0.40819600224494934
step: 40, loss: 0.22784654796123505
step: 50, loss: 0.3499963581562042
step: 60, loss: 0.4272569715976715
step: 70, loss: 0.5033296942710876
step: 80, loss: 0.18688811361789703
```

```
step: 90, loss: 0.28168007731437683
step: 100, loss: 0.4340846538543701
step: 110, loss: 0.24368229508399963
step: 120, loss: 0.3665584921836853
step: 130, loss: 0.31893762946128845
step: 140, loss: 0.1690794974565506
step: 150, loss: 0.3036502003669739
step: 160, loss: 0.22899647057056427
step: 170, loss: 0.3501497209072113
step: 180, loss: 0.1428123414516449
step: 190, loss: 0.11156661808490753
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.44406196213425125, f1=0.45454545454545453,
best_f1=0.45454545454545453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3518383502960205
step: 10, loss: 0.21790501475334167
step: 20, loss: 0.7958233952522278
step: 30, loss: 0.2302573174238205
step: 40, loss: 0.5265429019927979
step: 50, loss: 0.3559029996395111
step: 60, loss: 0.312531441450119
step: 70, loss: 0.2574595510959625
step: 80, loss: 0.07154521346092224
step: 90, loss: 0.23090335726737976
step: 100, loss: 0.10368657112121582
step: 110, loss: 0.18404659628868103
step: 120, loss: 0.15190556645393372
step: 130, loss: 0.2685992419719696
step: 140, loss: 0.25775450468063354
step: 150, loss: 0.25687506794929504
step: 160, loss: 0.23522980511188507
step: 170, loss: 0.07636209577322006
step: 180, loss: 0.04177209734916687
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.07930614799261093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.7267441860465116, f1=0.7548209366391184,
best_f1=0.7548209366391184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1822226196527481
step: 10, loss: 0.20970803499221802
step: 20, loss: 0.13068997859954834
step: 30, loss: 0.06775601208209991
step: 40, loss: 0.12579670548439026
step: 50, loss: 0.33022305369377136
step: 60, loss: 0.0768992006778717
step: 70, loss: 0.21839845180511475
step: 80, loss: 0.09865834563970566
step: 90, loss: 0.10903769731521606
step: 100, loss: 0.0906122624874115
step: 110, loss: 0.202290877699852
step: 120, loss: 0.22499848902225494
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 130, loss: 0.09629740566015244
step: 140, loss: 0.0723971426486969
step: 150, loss: 0.18076376616954803
step: 160, loss: 0.08253806829452515
step: 170, loss: 0.36993393301963806
step: 180, loss: 0.156705841422081
step: 190, loss: 0.06895384192466736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7848101265822786, f1=0.7989556135770235,
best_f1=0.7989556135770235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.054574768990278244
step: 10, loss: 0.054971758276224136
step: 20, loss: 0.04494563117623329
step: 30, loss: 0.0355040617287159
step: 40, loss: 0.11356059461832047
step: 50, loss: 0.2083781361579895
step: 60, loss: 0.06675877422094345

```
step: 70, loss: 0.055267613381147385
step: 80, loss: 0.026578886434435844
step: 90, loss: 0.059312693774700165
step: 100, loss: 0.12855122983455658
step: 110, loss: 0.2426275908946991
step: 120, loss: 0.05092446133494377
step: 130, loss: 0.04241310432553291
step: 140, loss: 0.1355665773153305
step: 150, loss: 0.025798315182328224
step: 160, loss: 0.022775152698159218
step: 170, loss: 0.050136275589466095
step: 180, loss: 0.041671112179756165
step: 190, loss: 0.03554225713014603
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8238482384823848, f1=0.8152173913043478,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10899973660707474
step: 10, loss: 0.04184728488326073
step: 20, loss: 0.010643244720995426
step: 30, loss: 0.02854117378592491
step: 40, loss: 0.0337098129093647
step: 50, loss: 0.03230859711766243
step: 60, loss: 0.01579730212688446
step: 70, loss: 0.14563632011413574
step: 80, loss: 0.04846258834004402
step: 90, loss: 0.09593629091978073
step: 100, loss: 0.18605844676494598
step: 110, loss: 0.2098902314901352
step: 120, loss: 0.07642994076013565
step: 130, loss: 0.31061145663261414
step: 140, loss: 0.07779417186975479
step: 150, loss: 0.04213150590658188
step: 160, loss: 0.05651584267616272
step: 170, loss: 0.015488862991333008
step: 180, loss: 0.03453710675239563
step: 190, loss: 0.04429464042186737
/usr/local/lib/python3.7/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=0.8067226890756303, f1=0.7774647887323942,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08158479630947113
step: 10, loss: 0.02459394931793213
step: 20, loss: 0.0080645419657230338
step: 30, loss: 0.056652218103408813
step: 40, loss: 0.026679061353206635
step: 50, loss: 0.06494564563035965
step: 60, loss: 0.01801997981965542
step: 70, loss: 0.05262523889541626
step: 80, loss: 0.04814421758055687
step: 90, loss: 0.01052758377045393
step: 100, loss: 0.24862614274024963
step: 110, loss: 0.013910523615777493
step: 120, loss: 0.016143420711159706
step: 130, loss: 0.04478849098086357
step: 140, loss: 0.016853593289852142
step: 150, loss: 0.02010355517268181
step: 160, loss: 0.23998332023620605
step: 170, loss: 0.11735478788614273
step: 180, loss: 0.028383633121848106
step: 190, loss: 0.07286633551120758
/usr/local/lib/python3.7/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=0.8210526315789474, f1=0.8072916666666666,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018525846302509308
step: 10, loss: 0.003305017249658704
step: 20, loss: 0.002476537600159645
step: 30, loss: 0.03423374146223068
step: 40, loss: 0.02471672184765339
step: 50, loss: 0.012704797089099884
step: 60, loss: 0.028781404718756676
step: 70, loss: 0.007489026989787817
step: 80, loss: 0.006579980719834566
step: 90, loss: 0.050479669123888016
step: 100, loss: 0.12569653987884521
step: 110, loss: 0.2708841562271118
step: 120, loss: 0.05172068625688553
step: 130, loss: 0.0258039440959692
step: 140, loss: 0.02873682975769043
```

```
step: 150, loss: 0.02933981455862522
step: 160, loss: 0.25339552760124207
step: 170, loss: 0.03357665613293648
step: 180, loss: 0.013071293942630291
step: 190, loss: 0.026623612269759178
/usr/local/lib/python3.7/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=0.8044692737430168, f1=0.8100558659217876,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13952067494392395
step: 10, loss: 0.03166124224662781
step: 20, loss: 0.011243084445595741
step: 30, loss: 0.05476123094558716
step: 40, loss: 0.020644132047891617
step: 50, loss: 0.07417488098144531
step: 60, loss: 0.26991233229637146
step: 70, loss: 0.0019677437376230955
step: 80, loss: 0.27609923481941223
step: 90, loss: 0.16178077459335327
step: 100, loss: 0.0038956052158027887
step: 110, loss: 0.02534080669283867
step: 120, loss: 0.00562829477712512
step: 130, loss: 0.0038443508092314005
step: 140, loss: 0.004490150138735771
step: 150, loss: 0.06651002913713455
step: 160, loss: 0.003906573634594679
step: 170, loss: 0.04168255627155304
step: 180, loss: 0.04924982041120529
step: 190, loss: 0.0074288188479840755
/usr/local/lib/python3.7/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=0.8238482384823848, f1=0.8333333333333334,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021626478992402554
step: 10, loss: 0.0011564777232706547
step: 20, loss: 0.025290701538324356
step: 30, loss: 0.0059397052973508835
step: 40, loss: 0.10423543304204941
step: 50, loss: 0.07624995708465576
step: 60, loss: 0.04183628037571907
```

```
step: 70, loss: 0.06244777888059616
step: 80, loss: 0.006746795028448105
step: 90, loss: 0.030005548149347305
step: 100, loss: 0.0513576865196228
step: 110, loss: 0.0008898738888092339
step: 120, loss: 0.011649219319224358
step: 130, loss: 0.0025322670117020607
step: 140, loss: 0.03950921818614006
step: 150, loss: 0.0049865045584738255
step: 160, loss: 0.01717929169535637
step: 170, loss: 0.03278333693742752
step: 180, loss: 0.04123859852552414
step: 190, loss: 0.001318547991104424
/usr/local/lib/python3.7/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=0.8222811671087534, f1=0.8225806451612903,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005108047043904662
step: 10, loss: 0.0013555458281189203
step: 20, loss: 0.0023025437258183956
step: 30, loss: 0.014133870601654053
step: 40, loss: 0.0018200414488092065
step: 50, loss: 0.0033500781282782555
step: 60, loss: 0.01803230494260788
step: 70, loss: 0.0009892110247164965
step: 80, loss: 0.006116189062595367
step: 90, loss: 0.003011920489370823
step: 100, loss: 0.0067242891527712345
step: 110, loss: 0.0017391839064657688
step: 120, loss: 0.010770969092845917
step: 130, loss: 0.006033570040017366
step: 140, loss: 0.10651182383298874
step: 150, loss: 0.0036535339895635843
step: 160, loss: 0.0007852168055251241
step: 170, loss: 0.002132249530404806
step: 180, loss: 0.014213286340236664
step: 190, loss: 0.013777642510831356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.805, f1=0.8, best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.002577533246949315
step: 10, loss: 0.0013688491890206933
step: 20, loss: 0.003641085932031274
step: 30, loss: 0.003253740957006812
step: 40, loss: 0.0005084297154098749
step: 50, loss: 0.0004474409797694534
step: 60, loss: 0.016674695536494255
step: 70, loss: 0.002147010527551174
step: 80, loss: 0.001643342082388699
step: 90, loss: 0.0014187361812219024
step: 100, loss: 0.014014320448040962
step: 110, loss: 0.02767685055732727
step: 120, loss: 0.0008585925679653883
step: 130, loss: 0.00933761615306139
step: 140, loss: 0.0008156928233802319
step: 150, loss: 0.004386023618280888
step: 160, loss: 0.0008994719828478992
step: 170, loss: 0.0012965891510248184
step: 180, loss: 0.007212154101580381
step: 190, loss: 0.0019363549072295427
/usr/local/lib/python3.7/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=0.8169761273209548, f1=0.8169761273209548,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017299780156463385
step: 10, loss: 0.032635752111673355
step: 20, loss: 0.0015934319235384464
step: 30, loss: 0.00650505255907774
step: 40, loss: 0.006017758511006832
step: 50, loss: 0.0007714694947935641
step: 60, loss: 0.03954511508345604
step: 70, loss: 0.002484223572537303
step: 80, loss: 0.030672717839479446
step: 90, loss: 0.0029970367904752493
step: 100, loss: 0.04747796431183815
step: 110, loss: 0.24369965493679047
step: 120, loss: 0.032564010471105576
step: 130, loss: 0.005484297405928373
step: 140, loss: 0.00200931285507977
step: 150, loss: 0.00627773767337203
step: 160, loss: 0.002433418296277523
step: 170, loss: 0.012579538859426975
step: 180, loss: 0.0009827121393755078
step: 190, loss: 0.04951115697622299
```

```
/usr/local/lib/python3.7/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=0.8067226890756303, f1=0.8056338028169014,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2205522209405899
step: 10, loss: 0.0009421881404705346
step: 20, loss: 0.04062305763363838
step: 30, loss: 0.052714634686670845
step: 40, loss: 0.004074919503182173
step: 50, loss: 0.0032286131754517555
step: 60, loss: 0.07608693093061447
step: 70, loss: 0.03991718590259552
step: 80, loss: 0.0027281460352241993
step: 90, loss: 0.011840438470244408
step: 100, loss: 0.004044839181005955
step: 110, loss: 0.0023893348407000303
step: 120, loss: 0.0013669586041942239
step: 130, loss: 0.0006242856616154313
step: 140, loss: 0.005552453920245171
step: 150, loss: 0.0016753539675846696
step: 160, loss: 0.0013301598373800516
step: 170, loss: 0.006316785234957933
step: 180, loss: 0.00278711155988276
step: 190, loss: 0.011789653450250626
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.8184281842818427,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004394523333758116
step: 10, loss: 0.022367611527442932
step: 20, loss: 0.0009817234240472317
step: 30, loss: 0.004637775011360645
step: 40, loss: 0.0011164878960698843
step: 50, loss: 0.0010113946627825499
step: 60, loss: 0.003910255618393421
step: 70, loss: 0.0006968051893636584
step: 80, loss: 0.008477194234728813
step: 90, loss: 0.00050976994680240750
step: 100, loss: 0.0005935318185947835
step: 110, loss: 0.0038392054848372936
```

```
step: 120, loss: 0.0031968082766979933
step: 130, loss: 0.0006980886682868004
step: 140, loss: 0.01807517744600773
step: 150, loss: 0.002138921758159995
step: 160, loss: 0.0015705717960372567
step: 170, loss: 0.0011639383155852556
step: 180, loss: 0.0015830559423193336
step: 190, loss: 0.0008935308433137834
/usr/local/lib/python3.7/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=0.8103896103896104, f1=0.8144329896907216,
best_f1=0.8152173913043478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010730773210525513
step: 10, loss: 0.03190881386399269
step: 20, loss: 0.10753772407770157
step: 30, loss: 0.02507386915385723
step: 40, loss: 0.002546961884945631
step: 50, loss: 0.0035258724819868803
step: 60, loss: 0.0020726113580167294
step: 70, loss: 0.0048629362136125565
step: 80, loss: 0.0011783677618950605
step: 90, loss: 0.005952172912657261
step: 100, loss: 0.0009885517647489905
step: 110, loss: 0.007077065762132406
step: 120, loss: 0.0014153467491269112
step: 130, loss: 0.0009504579356871545
step: 140, loss: 0.0014641679590567946
step: 150, loss: 0.001032872125506401
step: 160, loss: 0.0007046792306937277
step: 170, loss: 0.0011585387401282787
step: 180, loss: 0.0013992551248520613
step: 190, loss: 0.004804421216249466
/usr/local/lib/python3.7/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=0.8073878627968338, f1=0.8063660477453581,
best_f1=0.8152173913043478
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 [00:13, 157.19it/s]
load_f1 = 0.8091168091168092
real_f1 = 0.7734806629834254
733it [00:00, 3525.66it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:29, 147.03it/s]
```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

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

```
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general  \
  --dk product \
  --summarize \
  --da entry_swap  \
  --da attr_del


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

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

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

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

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4873254895210266
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5054770708084106
step: 20, loss: 0.31690728664398193
step: 30, loss: 0.4098995327949524
step: 40, loss: 0.5052679181098938
step: 50, loss: 0.32401180267333984
step: 60, loss: 0.5365761518478394
step: 70, loss: 0.30875512957572937
step: 80, loss: 0.2254856824874878
step: 90, loss: 0.2583373486995697
step: 100, loss: 0.15581606328487396
step: 110, loss: 0.35848790407180786
step: 120, loss: 0.3018357753753662
step: 130, loss: 0.32627081871032715
step: 140, loss: 0.4229556918144226
step: 150, loss: 0.2973198890686035
step: 160, loss: 0.3965550363063812
step: 170, loss: 0.30990856885910034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34674808382987976
step: 10, loss: 0.47264912724494934
step: 20, loss: 0.31793755292892456
step: 30, loss: 0.324687659740448
step: 40, loss: 0.09619180858135223

```
step: 50, loss: 0.43803641200065613
step: 60, loss: 0.1923772394657135
step: 70, loss: 0.5055564045906067
step: 80, loss: 0.23278355598449707
step: 90, loss: 0.2535772919654846
step: 100, loss: 0.4967910349369049
step: 110, loss: 0.2729383111000061
step: 120, loss: 0.23504863679409027
step: 130, loss: 0.5447263717651367
step: 140, loss: 0.5029798150062561
step: 150, loss: 0.43194150924682617
step: 160, loss: 0.46100538969039917
step: 170, loss: 0.3807348906993866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.23776223776223776, f1=0.24282560706401768,
best_f1=0.24282560706401768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6146436929702759
step: 10, loss: 0.2617599368095398
step: 20, loss: 0.2400262951850891
step: 30, loss: 0.2573564648628235
step: 40, loss: 0.39342737197875977
step: 50, loss: 0.5167526006698608
step: 60, loss: 0.2699207067489624
step: 70, loss: 0.25022387504577637
step: 80, loss: 0.4137437045574188
step: 90, loss: 0.5190675854682922
step: 100, loss: 0.29507580399513245
step: 110, loss: 0.16562460362911224
step: 120, loss: 0.5494199991226196
step: 130, loss: 0.4864661693572998
step: 140, loss: 0.4459221363067627
step: 150, loss: 0.2072954773902893
step: 160, loss: 0.1646500676870346
step: 170, loss: 0.3051157295703888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24772036474164133, f1=0.2392909896602659,
best_f1=0.2392909896602659
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35465624928474426
step: 10, loss: 0.5008967518806458
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.2109796404838562
step: 30, loss: 0.3741821050643921
step: 40, loss: 0.22306033968925476
step: 50, loss: 0.347097784280777
step: 60, loss: 0.6315826177597046
step: 70, loss: 0.29515978693962097
step: 80, loss: 0.49512234330177307
step: 90, loss: 0.3013450503349304
step: 100, loss: 0.4159686863422394
step: 110, loss: 0.4492860436439514
step: 120, loss: 0.4576859474182129
step: 130, loss: 0.33488285541534424
step: 140, loss: 0.2545034885406494
step: 150, loss: 0.6860480904579163
step: 160, loss: 0.13557352125644684
step: 170, loss: 0.24420003592967987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.25812115891132575, f1=0.24977856510186006,
best_f1=0.24977856510186006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34600648283958435
step: 10, loss: 0.34267982840538025
step: 20, loss: 0.3095529079437256
step: 30, loss: 0.29113736748695374
step: 40, loss: 0.2592916786670685
step: 50, loss: 0.24284721910953522
step: 60, loss: 0.26326870918273926
step: 70, loss: 0.3422408401966095
step: 80, loss: 0.07679253816604614
step: 90, loss: 0.5827853083610535

```
step: 100, loss: 0.2561170756816864
step: 110, loss: 0.25115275382995605
step: 120, loss: 0.13296879827976227
step: 130, loss: 0.25751012563705444
step: 140, loss: 0.18032661080360413
step: 150, loss: 0.3068154454231262
step: 160, loss: 0.25093767046928406
step: 170, loss: 0.3776320219039917
/usr/local/lib/python3.7/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=0.19433962264150942, f1=0.19412878787878787,
best_f1=0.24977856510186006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3090081214904785
step: 10, loss: 0.10766283422708511
step: 20, loss: 0.3808993101119995
step: 30, loss: 0.24640607833862305
step: 40, loss: 0.43944668769836426
step: 50, loss: 0.26750993728637695
step: 60, loss: 0.3872753381729126
step: 70, loss: 0.379086434841156
step: 80, loss: 0.1940499246120453
step: 90, loss: 0.3126520812511444
step: 100, loss: 0.17281773686408997
step: 110, loss: 0.4384777843952179
step: 120, loss: 0.4514329135417938
step: 130, loss: 0.5023095011711121
step: 140, loss: 0.23043577373027802
step: 150, loss: 0.17731711268424988
step: 160, loss: 0.323341965675354
step: 170, loss: 0.21546658873558044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2662993572084481, f1=0.24747937671860679,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2265133410692215
step: 10, loss: 0.4419078528881073
```

```
step: 20, loss: 0.4285511374473572
step: 30, loss: 0.45046013593673706
step: 40, loss: 0.1329195350408554
step: 50, loss: 0.4044762849807739
step: 60, loss: 0.504826009273529
step: 70, loss: 0.3540124297142029
step: 80, loss: 0.18020421266555786
step: 90, loss: 0.5061092376708984
step: 100, loss: 0.2684040367603302
step: 110, loss: 0.3096977174282074
step: 120, loss: 0.39261600375175476
step: 130, loss: 0.20469322800636292
step: 140, loss: 0.15021716058254242
step: 150, loss: 0.26968255639076233
step: 160, loss: 0.3388011157512665
step: 170, loss: 0.3340340852737427
/usr/local/lib/python3.7/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=0.24797047970479708, f1=0.23573573573573572,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4270770251750946
step: 10, loss: 0.5349050760269165
step: 20, loss: 0.20408260822296143
step: 30, loss: 0.2713419198989868
step: 40, loss: 0.23718321323394775
step: 50, loss: 0.3431341052055359
step: 60, loss: 0.38949379324913025
step: 70, loss: 0.2057914286851883
step: 80, loss: 0.2904226779937744
step: 90, loss: 0.5396417379379272
step: 100, loss: 0.24379104375839233
step: 110, loss: 0.46327289938926697
step: 120, loss: 0.31538480520248413
step: 130, loss: 0.28960278630256653
step: 140, loss: 0.45653778314590454
step: 150, loss: 0.22136038541793823
step: 160, loss: 0.1466793268918991
step: 170, loss: 0.40369483828544617
/usr/local/lib/python3.7/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=0.2584784601283226, f1=0.24817518248175183,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30438074469566345
step: 10, loss: 0.5739246010780334
step: 20, loss: 0.3194389045238495
step: 30, loss: 0.17172403633594513
step: 40, loss: 0.4141640067100525
step: 50, loss: 0.21771565079689026
step: 60, loss: 0.3455096185207367
step: 70, loss: 0.3608579635620117
step: 80, loss: 0.13814693689346313
step: 90, loss: 0.4282756447792053
step: 100, loss: 0.47323697805404663
step: 110, loss: 0.3435908854007721
step: 120, loss: 0.38491740822792053
step: 130, loss: 0.311524897813797
step: 140, loss: 0.3756285607814789
step: 150, loss: 0.3230048716068268
step: 160, loss: 0.22194252908229828
step: 170, loss: 0.36905840039253235
/usr/local/lib/python3.7/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=0.2574074074074074, f1=0.24861878453038674,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4190543293952942
step: 10, loss: 0.29928943514823914
step: 20, loss: 0.2569277882575989
step: 30, loss: 0.6130529642105103
step: 40, loss: 0.3923613429069519
step: 50, loss: 0.3878547251224518
step: 60, loss: 0.5316621661186218
step: 70, loss: 0.2702849805355072
step: 80, loss: 0.28319984674453735
step: 90, loss: 0.1930142641067505
step: 100, loss: 0.25300103425979614
step: 110, loss: 0.3102971613407135
step: 120, loss: 0.28149089217185974
step: 130, loss: 0.16969381272792816
step: 140, loss: 0.26960551738739014
step: 150, loss: 0.2744416296482086
step: 160, loss: 0.20889054238796234
step: 170, loss: 0.449588418006897
/usr/local/lib/python3.7/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=0.2589928057553957, f1=0.24329159212880141,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3673100173473358
step: 10, loss: 0.26571372151374817
step: 20, loss: 0.16524413228034973
step: 30, loss: 0.4621891975402832
step: 40, loss: 0.21935787796974182
step: 50, loss: 0.20161055028438568
step: 60, loss: 0.3257788419723511
step: 70, loss: 0.2997855544090271
step: 80, loss: 0.20250333845615387
step: 90, loss: 0.3539881706237793
step: 100, loss: 0.5043378472328186
step: 110, loss: 0.3483578860759735
step: 120, loss: 0.3560514450073242
step: 130, loss: 0.41319817304611206
step: 140, loss: 0.5506129264831543
step: 150, loss: 0.3379916250705719
step: 160, loss: 0.26223224401474
step: 170, loss: 0.16613034904003143
/usr/local/lib/python3.7/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=0.259047619047619, f1=0.25190839694656486,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3654576241970062
step: 10, loss: 0.2688542902469635
step: 20, loss: 0.23524293303489685
step: 30, loss: 0.27860718965530396
step: 40, loss: 0.3001883924007416
step: 50, loss: 0.3148192763328552
step: 60, loss: 0.3193127512931824
step: 70, loss: 0.16573096811771393
step: 80, loss: 0.22547602653503418
step: 90, loss: 0.28277385234832764
step: 100, loss: 0.2489178329706192
step: 110, loss: 0.4088343381881714
step: 120, loss: 0.28975626826286316
step: 130, loss: 0.3748728036880493
step: 140, loss: 0.2891676425933838
step: 150, loss: 0.2654683589935303
```

```
step: 160, loss: 0.6680076718330383
step: 170, loss: 0.36883872747421265
/usr/local/lib/python3.7/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=0.257372654155496, f1=0.24446412754650135,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3128098249435425
step: 10, loss: 0.2223292738199234
step: 20, loss: 0.18200217187404633
step: 30, loss: 0.3713805079460144
step: 40, loss: 0.4090500771999359
step: 50, loss: 0.23097974061965942
step: 60, loss: 0.22588007152080536
step: 70, loss: 0.32997873425483704
step: 80, loss: 0.19977208971977234
step: 90, loss: 0.29551881551742554
step: 100, loss: 0.17479759454727173
step: 110, loss: 0.43745699524879456
step: 120, loss: 0.19995194673538208
step: 130, loss: 0.27473464608192444
step: 140, loss: 0.5687770843505859
step: 150, loss: 0.3883131742477417
step: 160, loss: 0.3217271864414215
step: 170, loss: 0.286283940076828
/usr/local/lib/python3.7/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=0.2619254119687771, f1=0.24326672458731535,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3588055372238159
step: 10, loss: 0.2502058148384094
step: 20, loss: 0.3398469388484955
step: 30, loss: 0.40075385570526123
step: 40, loss: 0.21344323456287384
step: 50, loss: 0.2529812157154083
step: 60, loss: 0.3698819577693939
step: 70, loss: 0.3018244504928589
step: 80, loss: 0.16118092834949493
step: 90, loss: 0.44356822967529297
step: 100, loss: 0.40234193205833435
step: 110, loss: 0.2156704068183899
```

```
step: 120, loss: 0.2655576765537262
step: 130, loss: 0.39238241314888
step: 140, loss: 0.3156713843345642
step: 150, loss: 0.25541990995407104
step: 160, loss: 0.25878387689590454
step: 170, loss: 0.24481448531150818
/usr/local/lib/python3.7/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=0.25844594594594594, f1=0.24170212765957447,
best_f1=0.24747937671860679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11634942889213562
step: 10, loss: 0.390946626663208
step: 20, loss: 0.345400333404541
step: 30, loss: 0.5178542137145996
step: 40, loss: 0.16980937123298645
step: 50, loss: 0.19072610139846802
step: 60, loss: 0.32368946075439453
step: 70, loss: 0.17923903465270996
step: 80, loss: 0.2717546820640564
step: 90, loss: 0.2805939316749573
step: 100, loss: 0.46041321754455566
step: 110, loss: 0.49243929982185364
step: 120, loss: 0.3290283977985382
step: 130, loss: 0.4413702189922333
step: 140, loss: 0.33358126878738403
step: 150, loss: 0.1891280710697174
step: 160, loss: 0.20415280759334564
step: 170, loss: 0.2534081041812897
/usr/local/lib/python3.7/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=0.2611683848797251, f1=0.23924268502581755,
best_f1=0.24747937671860679
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, 237.49it/s]
load_f1 = 0.26371511068334935
real_f1 = 0.26201696512723843
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 146.47it/s]
```

4.3   DITTO DIRTY

4.3.1   DBLP-ACM - Running the matcher

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

```
  --dk product \
  --summarize \
  --da entry_swap  \
  --da attr_del


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

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

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

Defaults for this optimization level are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5996264219284058
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```
step: 10, loss: 0.41423115134239197
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.49187374114990234
step: 30, loss: 0.32953014969825745
step: 40, loss: 0.33340761065483093
step: 50, loss: 0.5054224729537964
step: 60, loss: 0.14054423570632935
step: 70, loss: 0.0697629302740097
step: 80, loss: 0.043936699628829956
step: 90, loss: 0.112335205078125
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.1772986352443695
step: 110, loss: 0.028069674968719482
step: 120, loss: 0.024417469277977943
step: 130, loss: 0.04578765481710434
step: 140, loss: 0.009730476886034012
step: 150, loss: 0.11440468579530716
step: 160, loss: 0.0021689916029572487
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.056355204433202744
step: 180, loss: 0.1933956891298294
step: 190, loss: 0.054573219269514084
step: 200, loss: 0.06437195092439651
step: 210, loss: 0.09494141489267349
step: 220, loss: 0.11004330217838287
step: 230, loss: 0.004860273562371731
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.964835164835165, f1=0.9655172413793103,
best_f1=0.9655172413793103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026608966290950775
step: 10, loss: 0.1204771175980568
step: 20, loss: 0.0063535673543810844
step: 30, loss: 0.011422688141465187
step: 40, loss: 0.07490997016429901
step: 50, loss: 0.00838110689073801
step: 60, loss: 0.08041112124919891
step: 70, loss: 0.0158731359988451
step: 80, loss: 0.0027350252494215965
step: 90, loss: 0.02719677798449993
```

```
step: 100, loss: 0.001495618256740272
step: 110, loss: 0.04304562136530876
step: 120, loss: 0.18009053170681
step: 130, loss: 0.014861221425235271
step: 140, loss: 0.0016427512746304274
step: 150, loss: 0.055541280657052994
step: 160, loss: 0.018008721992373466
step: 170, loss: 0.003499654121696949
step: 180, loss: 0.003143591107800603
step: 190, loss: 0.11730612814426422
step: 200, loss: 0.011483917012810707
step: 210, loss: 0.03976481035351753
step: 220, loss: 0.0028133345767855644
step: 230, loss: 0.003112497041001916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9743589743589743, f1=0.9684684684684683,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007017180789262056
step: 10, loss: 0.002858229447156191
step: 20, loss: 0.004943021573126316
step: 30, loss: 0.01657206192612648
step: 40, loss: 0.004377441946417093
step: 50, loss: 0.018725188449025154
step: 60, loss: 0.007703571114689112
step: 70, loss: 0.006520008202642202
step: 80, loss: 0.05562063679099083
step: 90, loss: 0.007463347632437944
step: 100, loss: 0.004310350399464369
step: 110, loss: 0.01153668574988842
step: 120, loss: 0.0007300458382815123
step: 130, loss: 0.042178280651569366
step: 140, loss: 0.0008323206566274166
step: 150, loss: 0.1549626737833023
step: 160, loss: 0.0019086444517597556
step: 170, loss: 0.006151252426207066
step: 180, loss: 0.05890418961644173
step: 190, loss: 0.008630158379673958
step: 200, loss: 0.013477870263159275
step: 210, loss: 0.0013212591875344515
```

```
step: 220, loss: 0.08262075483798981
step: 230, loss: 0.031388185918331146
/usr/local/lib/python3.7/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=0.9675977653631285, f1=0.963963963963964,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007463029120117426
step: 10, loss: 0.0007467474206350744
step: 20, loss: 0.007710066623985767
step: 30, loss: 0.0004882319481112063
step: 40, loss: 0.0038357949815690517
step: 50, loss: 0.00881136953830719
step: 60, loss: 0.0013304605381563306
step: 70, loss: 0.06329673528671265
step: 80, loss: 0.19241778552532196
step: 90, loss: 0.034891191869974136
step: 100, loss: 0.05030504986643791
step: 110, loss: 0.003356514498591423
step: 120, loss: 0.006389433518052101
step: 130, loss: 0.015777887776494026
step: 140, loss: 0.03146769478917122
step: 150, loss: 0.012226917780935764
step: 160, loss: 0.05068875849246979
step: 170, loss: 0.0017586752073839307
step: 180, loss: 0.1344553828239441
step: 190, loss: 0.001338987029157579
step: 200, loss: 0.06735499203205109
step: 210, loss: 0.010178954340517521
step: 220, loss: 0.001180286519229412
step: 230, loss: 0.00887903943657875
/usr/local/lib/python3.7/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=0.9720670391061451, f1=0.9742441209406495,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01841363124549389
step: 10, loss: 0.0014327788958325982
step: 20, loss: 0.21833544969558716
step: 30, loss: 0.0017393659800291061
step: 40, loss: 0.004393387585878372
step: 50, loss: 0.0036587691865861416
```

```
step: 60, loss: 0.030506523326039314
step: 70, loss: 0.10815352201461792
step: 80, loss: 0.17324651777744293
step: 90, loss: 0.09361806511878967
step: 100, loss: 0.003544570878148079
step: 110, loss: 0.016994819045066833
step: 120, loss: 0.004082642961293459
step: 130, loss: 0.007077691610902548
step: 140, loss: 0.010495278052985668
step: 150, loss: 0.001353075378574431
step: 160, loss: 0.0021789581514894962
step: 170, loss: 0.042588066309690475
step: 180, loss: 0.006817749235779047
step: 190, loss: 0.1735951155424118
step: 200, loss: 0.11712320894002914
step: 210, loss: 0.011790779419243336
step: 220, loss: 0.02413678914308548
step: 230, loss: 0.02377081662416458
/usr/local/lib/python3.7/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=0.9711111111111111, f1=0.9698324022346367,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000622013583779335
step: 10, loss: 0.0007322783931158483
step: 20, loss: 0.0008534521912224591
step: 30, loss: 0.0006356746307574213
step: 40, loss: 0.0002093868824886158
step: 50, loss: 0.011881203390657902
step: 60, loss: 0.0029967748560011387
step: 70, loss: 0.15878456830978394
step: 80, loss: 0.001694048405624926
step: 90, loss: 0.023624788969755173
step: 100, loss: 0.0008063338464125991
step: 110, loss: 0.06882397830486298
step: 120, loss: 0.02029605209827423
step: 130, loss: 0.002755449153482914
step: 140, loss: 0.0035387284588068724
step: 150, loss: 0.00026299498858861625
step: 160, loss: 0.004188714548945427
step: 170, loss: 0.0011442301329225302
step: 180, loss: 0.051778942346572876
step: 190, loss: 0.013979103416204453
step: 200, loss: 0.013695552945137024
step: 210, loss: 0.0006266257842071354
```

```
step: 220, loss: 0.04429107531905174
step: 230, loss: 0.0003412202058825642
/usr/local/lib/python3.7/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=0.9742441209406495, f1=0.9708520179372198,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000763358490075916
step: 10, loss: 0.0013065087841823697
step: 20, loss: 0.0021348227746784687
step: 30, loss: 0.000480696267914027
step: 40, loss: 0.0005001242971047759
step: 50, loss: 0.00020496040815487504
step: 60, loss: 0.0048460578545928
step: 70, loss: 0.0002956320531666279
step: 80, loss: 0.00036551529774442315
step: 90, loss: 0.007724246475845575
step: 100, loss: 0.0008316994644701481
step: 110, loss: 0.00034949855762533844
step: 120, loss: 0.00024315621703863144
step: 130, loss: 0.0007386863580904901
step: 140, loss: 0.00015267737035173923
step: 150, loss: 0.012298932299017906
step: 160, loss: 0.00021307315910235047
step: 170, loss: 0.0005568789201788604
step: 180, loss: 0.0005775887402705848
step: 190, loss: 0.01844015344977379
step: 200, loss: 0.005165088456124067
step: 210, loss: 0.00083641015522602141
step: 220, loss: 0.0017796418396756053
step: 230, loss: 0.0031722260173410177
/usr/local/lib/python3.7/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=0.9706546275395034, f1=0.9695603156708005,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041793332172498107
step: 10, loss: 0.004778815433382988
step: 20, loss: 0.00023592614161316305
step: 30, loss: 0.00014472022303380072
step: 40, loss: 0.14791016280651093
step: 50, loss: 0.003464836394413057
```

```
step: 60, loss: 0.0004078398924320936
step: 70, loss: 9.921807941282168e-05
step: 80, loss: 0.007036715280264616
step: 90, loss: 0.0001671879435889423
step: 100, loss: 0.0014194403775036335
step: 110, loss: 0.004906294867396355
step: 120, loss: 0.0005906241131015122
step: 130, loss: 0.0053719934076070786
step: 140, loss: 0.00039193680277094245
step: 150, loss: 0.1272616684436798
step: 160, loss: 0.0008082397980615497
step: 170, loss: 0.0013517525512725115
step: 180, loss: 0.00044803626951761544
step: 190, loss: 0.03557208925485611
step: 200, loss: 0.009128118865191936
step: 210, loss: 0.028450697660446167
step: 220, loss: 0.0007663769065402448
step: 230, loss: 0.0008291574195027351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9755011135857461, 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: 0.0006186055252328515
step: 10, loss: 0.0018140280153602362
step: 20, loss: 0.0022390864323824644
step: 30, loss: 0.000632846262305975
step: 40, loss: 0.001463029533624649
step: 50, loss: 0.0006406793254427612
step: 60, loss: 0.0010245100129395723
step: 70, loss: 0.013200202025473118
step: 80, loss: 0.0002995917748194188
step: 90, loss: 0.03622157871723175
step: 100, loss: 0.00021655537420883775
step: 110, loss: 9.717075590742752e-05
step: 120, loss: 0.0019056509481742978
step: 130, loss: 0.02854163385927677
step: 140, loss: 0.0004473667941056192
step: 150, loss: 0.0002619618608150631
step: 160, loss: 0.00019569713913369924
step: 170, loss: 0.0322677306830883
```

```
step: 180, loss: 0.0011636497220024467
step: 190, loss: 0.00014159386046230793
step: 200, loss: 0.0077699352987110615
step: 210, loss: 0.03653217852115631
step: 220, loss: 0.0002919557737186551
step: 230, loss: 0.0003974877472501248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.978675645342312, f1=0.9785794813979707,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006000830908305943
step: 10, loss: 0.00025740166689322889
step: 20, loss: 0.01437954418361187
step: 30, loss: 0.00021693359303753823
step: 40, loss: 0.00156982545740090242
step: 50, loss: 6.342471169773489e-05
step: 60, loss: 0.00019174348562955856
step: 70, loss: 0.03160438314080238
step: 80, loss: 0.02396402694284916
step: 90, loss: 5.424641494755633e-05
step: 100, loss: 9.017287811730057e-05
step: 110, loss: 0.05044250190258026
step: 120, loss: 0.0002879848761949688
step: 130, loss: 0.00011816487676696852
step: 140, loss: 8.268943201983348e-05
step: 150, loss: 0.007844756357371807
step: 160, loss: 3.459447543718852e-05
step: 170, loss: 8.638693543616682e-05
step: 180, loss: 0.0063554770313203335
step: 190, loss: 0.0008846002165228128
step: 200, loss: 0.00019563693786039948
step: 210, loss: 0.004607517272233963
step: 220, loss: 0.008291302248835564
step: 230, loss: 6.363513966789469e-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 10: dev_f1=0.9808342728297633, f1=0.9751693002257337,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.910710049443878e-05
step: 10, loss: 0.00010818693408509716
step: 20, loss: 6.232200394151732e-05
step: 30, loss: 6.743764242855832e-05
step: 40, loss: 2.695062903512735e-05
step: 50, loss: 0.00010268910409649834
step: 60, loss: 0.015213178470730782
step: 70, loss: 6.038008359610103e-05
step: 80, loss: 0.00043641190859489143
step: 90, loss: 0.006539235357195139
step: 100, loss: 7.963446114445105e-05
step: 110, loss: 8.071122283581644e-05
step: 120, loss: 0.00020773580763489008
step: 130, loss: 0.00023682607570663095
step: 140, loss: 0.00017237986321561038
step: 150, loss: 0.0004122535465285182
step: 160, loss: 0.03900328651070595
step: 170, loss: 0.002988096559420228
step: 180, loss: 0.00029909590375609696
step: 190, loss: 0.0003993787686340511
step: 200, loss: 0.004009068477898836
step: 210, loss: 0.00027938219136558473
step: 220, loss: 0.001793142408132553
step: 230, loss: 0.0012225121026858687
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819819819819819, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.168225809233263e-05
step: 10, loss: 0.00016301879077218473
step: 20, loss: 0.0014651702949777246
step: 30, loss: 0.004428275860846043
step: 40, loss: 4.997446740162559e-05
step: 50, loss: 0.009632868692278862
```

```
step: 60, loss: 5.727443203795701e-05
step: 70, loss: 3.0427090678131208e-05
step: 80, loss: 2.204942393291276e-05
step: 90, loss: 0.00024259518249891698
step: 100, loss: 2.8959388146176934e-05
step: 110, loss: 2.223931187472772e-05
step: 120, loss: 6.209798448253423e-05
step: 130, loss: 3.0292907467810437e-05
step: 140, loss: 2.9808859835611656e-05
step: 150, loss: 0.00019086616521235555
step: 160, loss: 0.0018733515171334147
step: 170, loss: 0.015099437907338142
step: 180, loss: 3.294447378721088e-05
step: 190, loss: 5.399685323936865e-05
step: 200, loss: 3.401374488021247e-05
step: 210, loss: 7.142373942770064e-05
step: 220, loss: 0.0016536294715479016
step: 230, loss: 6.017094710841775e-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=0.9807909604519773, f1=0.9762174405436014,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.679631769657135e-05
step: 10, loss: 0.00017514685168862343
step: 20, loss: 3.2956304494291544e-05
step: 30, loss: 2.9612952857860364e-05
step: 40, loss: 3.798075340455398e-05
step: 50, loss: 0.001677719410508871
step: 60, loss: 2.502203824406024e-05
step: 70, loss: 0.002819816814735532
step: 80, loss: 0.000513655599206686
step: 90, loss: 0.00010654771176632494
step: 100, loss: 4.898957922705449e-05
step: 110, loss: 0.0016496197786182165
step: 120, loss: 0.0003964553470723331
step: 130, loss: 3.0505510949296877e-05
step: 140, loss: 6.28931520623155e-05
step: 150, loss: 3.564371581887826e-05
step: 160, loss: 0.0007986900163814425
step: 170, loss: 3.806288805208169e-05
step: 180, loss: 0.004162368830293417
step: 190, loss: 0.0005343465018086135
step: 200, loss: 4.1269700886914507e-05
step: 210, loss: 0.00790236797183752
```

```
step: 220, loss: 3.3775693736970425e-05
step: 230, loss: 2.749547638813965e-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=0.9819413092550789, f1=0.976271186440678,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.632919859024696e-05
step: 10, loss: 1.9520148271112703e-05
step: 20, loss: 0.00023495122215005234
step: 30, loss: 6.14161544945091e-05
step: 40, loss: 9.514481644146144e-05
step: 50, loss: 2.1710326109314337e-05
step: 60, loss: 0.0014936751686036587
step: 70, loss: 5.737107858294621e-05
step: 80, loss: 0.11112657934427261
step: 90, loss: 3.481072417343967e-05
step: 100, loss: 0.005170739255845547
step: 110, loss: 0.00810649711638689
step: 120, loss: 6.818571273470297e-05
step: 130, loss: 0.00010501484211999923
step: 140, loss: 9.779282117960975e-05
step: 150, loss: 2.0816300093429163e-05
step: 160, loss: 7.876657764427364e-05
step: 170, loss: 3.217763878637925e-05
step: 180, loss: 0.000723521108739078
step: 190, loss: 0.00010671485506463796
step: 200, loss: 0.00012013032392133027
step: 210, loss: 3.3379645174136385e-05
step: 220, loss: 0.00034652065369300544
step: 230, loss: 0.0014249020023271441
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9808773903262092,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0189181100577116
step: 10, loss: 2.434415728203021e-05
step: 20, loss: 0.0005116064567118883
step: 30, loss: 3.1320814741775393e-05
step: 40, loss: 2.1278196072671562e-05
step: 50, loss: 2.00118483917322e-05
```

```
step: 60, loss: 0.0015647505642846227
step: 70, loss: 0.0009102699696086347
step: 80, loss: 3.6937293771188706e-05
step: 90, loss: 3.140840635751374e-05
step: 100, loss: 1.6875299479579553e-05
step: 110, loss: 3.1633902835892513e-05
step: 120, loss: 0.06864258646965027
step: 130, loss: 0.00016762745508458465
step: 140, loss: 0.00012035746476612985
step: 150, loss: 0.00012840940325986594
step: 160, loss: 0.003021897515282035
step: 170, loss: 2.9129758331691846e-05
step: 180, loss: 3.03370561596239e-05
step: 190, loss: 0.0003142102505080402
step: 200, loss: 0.0019587459973990917
step: 210, loss: 0.000869921175763011
step: 220, loss: 2.266418050567154e-05
step: 230, loss: 3.4356562537141144e-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 15: dev_f1=0.9820224719101124, f1=0.9808773903262092,
best_f1=0.9808773903262092
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 [00:16, 149.22it/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 [00:32, 137.41it/s]
```

4.3.2 DBLP-GoogleScholar - Running the matcher

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


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

     # Step 3: run the matcher
     !CUDA_VISIBLE_DEVICES=0 python matcher.py \
       --task Dirty/DBLP-GoogleScholar \
```

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

Selected optimization level 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: 0.7186448574066162
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.460455060005188
step: 20, loss: 0.2690352201461792

```
step: 30, loss: 0.347439169883728
step: 40, loss: 0.36134013533592224
step: 50, loss: 0.5286909937858582
step: 60, loss: 0.2598314583301544
step: 70, loss: 0.30148181319236755
step: 80, loss: 0.10539203137159348
step: 90, loss: 0.17641660571098328
step: 100, loss: 0.35456305742263794
step: 110, loss: 0.13722427189350128
step: 120, loss: 0.15268810093402863
step: 130, loss: 0.10854081809520721
step: 140, loss: 0.2682097256183624
step: 150, loss: 0.022409385070204735
step: 160, loss: 0.140555739402771
step: 170, loss: 0.04879331588745117
step: 180, loss: 0.14429490268230438
step: 190, loss: 0.08909675478935242
step: 200, loss: 0.054541364312171936
step: 210, loss: 0.1022602990269661
step: 220, loss: 0.020344911143183708
step: 230, loss: 0.19426366686820984
step: 240, loss: 0.03743956238031387
step: 250, loss: 0.07271154969930649
step: 260, loss: 0.2741912007331848
step: 270, loss: 0.25591421127319336
step: 280, loss: 0.058123793452978134
step: 290, loss: 0.14704664051532745
step: 300, loss: 0.12488958984613419
step: 310, loss: 0.1172792911529541
step: 320, loss: 0.20650185644626617
step: 330, loss: 0.06164535880088806
step: 340, loss: 0.3024817407131195
step: 350, loss: 0.16426779329776764
step: 360, loss: 0.02775481343269348
step: 370, loss: 0.04230176657438278
step: 380, loss: 0.048786357045173645
step: 390, loss: 0.02475050278007984
step: 400, loss: 0.06943628937005997
step: 410, loss: 0.28680798411369324
step: 420, loss: 0.016440991312265396
step: 430, loss: 0.1949646770954132
step: 440, loss: 0.06412523239850998
step: 450, loss: 0.03153136372566223
step: 460, loss: 0.06761007010936737
step: 470, loss: 0.06700380146503448
step: 480, loss: 0.11418479681015015
step: 490, loss: 0.14032983779907227
step: 500, loss: 0.05565331503748894
```

```
step: 510, loss: 0.0544702373445034
step: 520, loss: 0.21439722180366516
step: 530, loss: 0.021546177566051483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9279486002753558, f1=0.9249658935879944,
best_f1=0.9249658935879944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02535095252096653
step: 10, loss: 0.05581538379192352
step: 20, loss: 0.06578513979911804
step: 30, loss: 0.022162307053804398
step: 40, loss: 0.05658867210149765
step: 50, loss: 0.19900719821453094
step: 60, loss: 0.054472703486680984
step: 70, loss: 0.043659426271915436
step: 80, loss: 0.043421678245067596
step: 90, loss: 0.015326114371418953
step: 100, loss: 0.13339939713478088
step: 110, loss: 0.011169811710715294
step: 120, loss: 0.042785923928022385
step: 130, loss: 0.008860527537763119
step: 140, loss: 0.04194476827979088
step: 150, loss: 0.06373754143714905
step: 160, loss: 0.04531366005539894
step: 170, loss: 0.057006582617759705
step: 180, loss: 0.09754649549722672
step: 190, loss: 0.10063633322715759
step: 200, loss: 0.08730008453130722
step: 210, loss: 0.05711200833320618
step: 220, loss: 0.001667427597567439
step: 230, loss: 0.04985044151544571
step: 240, loss: 0.06540986895561218
step: 250, loss: 0.017663966864347458
step: 260, loss: 0.23119288682937622
step: 270, loss: 0.06236805021762848
step: 280, loss: 0.0648108497262001
step: 290, loss: 0.06918967515230179
step: 300, loss: 0.044920455664396286
step: 310, loss: 0.11778229475021362
step: 320, loss: 0.10398536175489426
```

```
step: 330, loss: 0.07796904444694519
step: 340, loss: 0.12771548330783844
step: 350, loss: 0.005587911233305931
step: 360, loss: 0.04227720573544502
step: 370, loss: 0.02027716115117073
step: 380, loss: 0.2861660420894623
step: 390, loss: 0.015928559005260468
step: 400, loss: 0.14180301129817963
step: 410, loss: 0.016849856823682785
step: 420, loss: 0.039457108825445175
step: 430, loss: 0.08027518540620804
step: 440, loss: 0.004494306165724993
step: 450, loss: 0.01593746244907379
step: 460, loss: 0.016965417191386223
step: 470, loss: 0.017880715429782867
step: 480, loss: 0.03103390522301197
step: 490, loss: 0.0828116312623024
step: 500, loss: 0.05874289199709892
step: 510, loss: 0.12137927860021591
step: 520, loss: 0.3000112771987915
step: 530, loss: 0.29597747325897217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.935064935064935, f1=0.9393237610004631,
best_f1=0.9393237610004631
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11348091065883636
step: 10, loss: 0.0579320527613163
step: 20, loss: 0.020050346851348877
step: 30, loss: 0.04286544397473335
step: 40, loss: 0.15732401609420776
step: 50, loss: 0.046189043670892715
step: 60, loss: 0.10022515058517456
step: 70, loss: 0.011867406778037548
step: 80, loss: 0.24223557114601135
step: 90, loss: 0.019818559288978577
step: 100, loss: 0.04801405221223831
step: 110, loss: 0.04407436028122902
step: 120, loss: 0.043483756482601166
step: 130, loss: 0.0243443064391613
step: 140, loss: 0.009466899558901787
```

```
step: 150, loss: 0.16798654198646545
step: 160, loss: 0.029392391443252563
step: 170, loss: 0.014554517343640327
step: 180, loss: 0.010183432139456272
step: 190, loss: 0.018170148134231567
step: 200, loss: 0.016446713358163834
step: 210, loss: 0.013883179053664207
step: 220, loss: 0.20094507932662964
step: 230, loss: 0.04496181383728981
step: 240, loss: 0.03667246177792549
step: 250, loss: 0.030541053041815758
step: 260, loss: 0.1323421597480774
step: 270, loss: 0.009800927713513374
step: 280, loss: 0.0183465164154768
step: 290, loss: 0.0064355614595115185
step: 300, loss: 0.07090142369270325
step: 310, loss: 0.11209519952535629
step: 320, loss: 0.029021425172686577
step: 330, loss: 0.01923673413693905
step: 340, loss: 0.02318241074681282
step: 350, loss: 0.12240859121084213
step: 360, loss: 0.008298506028950214
step: 370, loss: 0.040218502283096313
step: 380, loss: 0.0024429995100945234
step: 390, loss: 0.13170436024665833
step: 400, loss: 0.06711240857839584
step: 410, loss: 0.0033766147680580616
step: 420, loss: 0.001558607560582459
step: 430, loss: 0.013633172027766705
step: 440, loss: 0.32292789220809937
step: 450, loss: 0.06764116138219833
step: 460, loss: 0.057571835815906525
step: 470, loss: 0.034189652651548386
step: 480, loss: 0.14041729271411896
step: 490, loss: 0.0022862800396978855
step: 500, loss: 0.010394240729510784
step: 510, loss: 0.012235410511493683
step: 520, loss: 0.0074867019429802895
step: 530, loss: 0.010843731462955475
/usr/local/lib/python3.7/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=0.934752429430819, f1=0.9340710004610421,
best_f1=0.9393237610004631
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01716964691877365
```

```
step: 10, loss: 0.013379527255892754
step: 20, loss: 0.139274001121521
step: 30, loss: 0.07296578586101532
step: 40, loss: 0.004009617492556572
step: 50, loss: 0.04837764799594879
step: 60, loss: 0.052441660314798355
step: 70, loss: 0.023360997438430786
step: 80, loss: 0.03422355651855469
step: 90, loss: 0.02733660489320755
step: 100, loss: 0.006607437506318092
step: 110, loss: 0.19076040387153625
step: 120, loss: 0.03660738095641136
step: 130, loss: 0.2594664692878723
step: 140, loss: 0.04052351787686348
step: 150, loss: 0.03445904701948166
step: 160, loss: 0.00422641821205616
step: 170, loss: 0.05421742796897888
step: 180, loss: 0.10033079981803894
step: 190, loss: 0.05657835677266121
step: 200, loss: 0.10879656672477722
step: 210, loss: 0.006175194401293993
step: 220, loss: 0.003959599416702986
step: 230, loss: 0.024067357182502747
step: 240, loss: 0.009373056702315807
step: 250, loss: 0.11232153326272964
step: 260, loss: 0.023190869018435478
step: 270, loss: 0.0912502184510231
step: 280, loss: 0.0051855044439435005
step: 290, loss: 0.016451839357614517
step: 300, loss: 0.00517575116828084
step: 310, loss: 0.05247694253921509
step: 320, loss: 0.14577844738960266
step: 330, loss: 0.025320764631032944
step: 340, loss: 0.030816754326224327
step: 350, loss: 0.13765937089920044
step: 360, loss: 0.026027293875813484
step: 370, loss: 0.00667763315141201
step: 380, loss: 0.005860021337866783
step: 390, loss: 0.0014456084463745356
step: 400, loss: 0.10932153463363647
step: 410, loss: 0.03515252470970154
step: 420, loss: 0.017375705763697624
step: 430, loss: 0.017251605167984962
step: 440, loss: 0.010108553804457188
step: 450, loss: 0.10748329013586044
step: 460, loss: 0.02481193095445633
step: 470, loss: 0.0023161321878433228
step: 480, loss: 0.024105068296194077
```

```
step: 490, loss: 0.012510527856647968
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 500, loss: 0.07766276597976685
step: 510, loss: 0.06917091459035873
step: 520, loss: 0.022394994273781776
step: 530, loss: 0.08892432600259781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9418334108887855, f1=0.9423887587822014,
best_f1=0.9423887587822014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004187923390418291
step: 10, loss: 0.008112482726573944
step: 20, loss: 0.012451455928385258
step: 30, loss: 0.009998179040849209
step: 40, loss: 0.03006070852279663
step: 50, loss: 0.09937641769647598
step: 60, loss: 0.012360069900751114
step: 70, loss: 0.040882714092731476
step: 80, loss: 0.0605006143450737
step: 90, loss: 0.22247762978076935
step: 100, loss: 0.14170221984386444
step: 110, loss: 0.02505306899547577
step: 120, loss: 0.17390811443328857
step: 130, loss: 0.005743448622524738
step: 140, loss: 0.1841253638267517
step: 150, loss: 0.044752635061740875
step: 160, loss: 0.0019698303658515215
step: 170, loss: 0.1986982673406601
step: 180, loss: 0.018109988421201706
step: 190, loss: 0.006997956428676844
step: 200, loss: 0.005863126367330551
step: 210, loss: 0.00625689746811986
step: 220, loss: 0.009228737093508244
step: 230, loss: 0.00893586128950119
step: 240, loss: 0.01234822254627943
step: 250, loss: 0.11895233392715454
step: 260, loss: 0.007985529489815235
step: 270, loss: 0.05864726006984711
step: 280, loss: 0.022366076707839966
step: 290, loss: 0.0120133887976408
```

```
step: 300, loss: 0.14822877943515778
step: 310, loss: 0.04617658257484436
step: 320, loss: 0.01691366918385029
step: 330, loss: 0.0008525456069037318
step: 340, loss: 0.005902957636862993
step: 350, loss: 0.00031753219082020223
step: 360, loss: 0.0003504588094074279
step: 370, loss: 0.000603717053309083
step: 380, loss: 0.001100044697523117
step: 390, loss: 0.023668425157666206
step: 400, loss: 0.004157899878919125
step: 410, loss: 0.04406479001045227
step: 420, loss: 0.21652290225028992
step: 430, loss: 0.12124375998973846
step: 440, loss: 0.0014039423549547791
step: 450, loss: 0.12216492742300034
step: 460, loss: 0.15864425897598267
step: 470, loss: 0.03296954929828644
step: 480, loss: 0.049155257642269135
step: 490, loss: 0.018788054585456848
step: 500, loss: 0.007463450543582439
step: 510, loss: 0.00283808377571404
step: 520, loss: 0.11288326978683472
step: 530, loss: 0.11065052449703217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9465437788018433, f1=0.9406858202038925,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03383733332157135
step: 10, loss: 0.1198991909623146
step: 20, loss: 0.00815179105848074
step: 30, loss: 0.0018200416816398501
step: 40, loss: 0.030251597985625267
step: 50, loss: 0.0014573431108146906
step: 60, loss: 0.07067698985338211
step: 70, loss: 0.06170589476823807
step: 80, loss: 0.001708430703729391
step: 90, loss: 0.011229590512812138
step: 100, loss: 0.007839974015951157
step: 110, loss: 0.005189025774598122
```

```
step: 120, loss: 0.02258143201470375
step: 130, loss: 0.002353216987103224
step: 140, loss: 0.0018129721283912659
step: 150, loss: 0.0066413311287760735
step: 160, loss: 0.013030187226831913
step: 170, loss: 0.0103166988119483
step: 180, loss: 0.0023162602446973324
step: 190, loss: 0.1738467961549759
step: 200, loss: 0.010971717536449432
step: 210, loss: 0.005613725166767836
step: 220, loss: 0.010349704883992672
step: 230, loss: 0.002370045054703951
step: 240, loss: 0.0017549630720168352
step: 250, loss: 0.026435302570462227
step: 260, loss: 0.004353450611233711
step: 270, loss: 0.0065102423541247845
step: 280, loss: 0.0027675663586705923
step: 290, loss: 0.0019456587033346295
step: 300, loss: 0.004625572822988033
step: 310, loss: 0.09526592493057251
step: 320, loss: 0.0003841579891741276
step: 330, loss: 0.004011860117316246
step: 340, loss: 0.0009319314267486334
step: 350, loss: 0.005938360467553139
step: 360, loss: 0.03606957197189331
step: 370, loss: 0.04714438319206238
step: 380, loss: 0.004363198298960924
step: 390, loss: 0.006568139884620905
step: 400, loss: 0.0005306490929797292
step: 410, loss: 0.009546950459480286
step: 420, loss: 0.01641113869845867
step: 430, loss: 0.0003902848984580487
step: 440, loss: 0.0007446618983522058
step: 450, loss: 0.1964641809463501
step: 460, loss: 0.0030140771996229887
step: 470, loss: 0.009236917831003666
step: 480, loss: 0.00781959667801857
step: 490, loss: 0.14526566863059998
step: 500, loss: 0.0018651693826541305
step: 510, loss: 0.07935388386249542
step: 520, loss: 0.003939149901270866
step: 530, loss: 0.01234086882323026
/usr/local/lib/python3.7/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=0.9439035697728326, f1=0.9424326833797586,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019755472894757986
step: 10, loss: 0.005668392404913902
step: 20, loss: 0.040684979408979416
step: 30, loss: 0.01359938271343708
step: 40, loss: 0.031599510461091995
step: 50, loss: 0.00806918740272522
step: 60, loss: 0.0008070169715210795
step: 70, loss: 0.030942857265472412
step: 80, loss: 0.00014574236411135644
step: 90, loss: 0.00011416900088079274
step: 100, loss: 0.09837020933628082
step: 110, loss: 0.001369517296552658
step: 120, loss: 0.08292316645383835
step: 130, loss: 0.0013069359119981527
step: 140, loss: 0.00033991364762187004
step: 150, loss: 0.01309119164943695
step: 160, loss: 0.0002111914218403399
step: 170, loss: 0.002622405532747507
step: 180, loss: 0.17722457647323608
step: 190, loss: 0.02254006266593933
step: 200, loss: 0.01025114394724369
step: 210, loss: 0.00031573098385706544
step: 220, loss: 0.0004947239649482071
step: 230, loss: 0.0005254666320979595
step: 240, loss: 0.35994985699653625
step: 250, loss: 0.013789774850010872
step: 260, loss: 0.024699484929442406
step: 270, loss: 0.0036806685384362936
step: 280, loss: 0.004625795409083366
step: 290, loss: 0.002337370067834854
step: 300, loss: 0.000890491355676204
step: 310, loss: 0.0007017640164121985
step: 320, loss: 0.1450101137161255
step: 330, loss: 0.02955346181988716
step: 340, loss: 0.005306612234562635
step: 350, loss: 0.007086934056133032
step: 360, loss: 0.014028133824467659
step: 370, loss: 0.01840789057314396
step: 380, loss: 0.018497996032238007
step: 390, loss: 0.003185085253790021
step: 400, loss: 0.059499211609363556
step: 410, loss: 0.0832735225558281
step: 420, loss: 0.006246720440685749
step: 430, loss: 0.006255879066884518
step: 440, loss: 0.0021287265699356794
step: 450, loss: 0.07046685367822647
```

```
step: 460, loss: 0.0031831625383347273
step: 470, loss: 0.051791414618492126
step: 480, loss: 0.002923812484368682
step: 490, loss: 0.001167205860838294
step: 500, loss: 0.0029705464839935303
step: 510, loss: 0.00193110725376755
step: 520, loss: 0.004364676773548126
step: 530, loss: 0.010945575311779976
/usr/local/lib/python3.7/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=0.945845004668534, f1=0.9398040130657956,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023647604510188103
step: 10, loss: 0.003339362796396017
step: 20, loss: 0.0016859524184837937
step: 30, loss: 0.03339097276329994
step: 40, loss: 0.0003404550370760262
step: 50, loss: 0.014246759936213493
step: 60, loss: 0.0006975653232075274
step: 70, loss: 0.002834787592291832
step: 80, loss: 0.0005082981660962105
step: 90, loss: 0.00033785036066547036
step: 100, loss: 0.0003643118543550372
step: 110, loss: 0.00038964481791481376
step: 120, loss: 0.0014094064245000482
step: 130, loss: 0.00030181446345523
step: 140, loss: 0.00041248759953305125
step: 150, loss: 0.00042179308366030455
step: 160, loss: 0.0013000750914216042
step: 170, loss: 0.06672590970993042
step: 180, loss: 0.022280121222138405
step: 190, loss: 0.022127319127321243
step: 200, loss: 0.08415520191192627
step: 210, loss: 0.2180906981229782
step: 220, loss: 0.0043655140325427055
step: 230, loss: 0.07038766145706177
step: 240, loss: 0.0042243460193276405
step: 250, loss: 0.0006961555918678641
step: 260, loss: 0.025159019976854324
step: 270, loss: 0.00036752992309629917
step: 280, loss: 0.00030239022453315556
step: 290, loss: 0.021891877055168152
step: 300, loss: 6.679342914139852e-05
step: 310, loss: 0.04448173567652702
```

```
step: 320, loss: 0.007020716555416584
step: 330, loss: 0.0006881188601255417
step: 340, loss: 0.023716865107417107
step: 350, loss: 0.002448529237881303
step: 360, loss: 0.03940483182668686
step: 370, loss: 0.0162163358181715
step: 380, loss: 0.0042526572942733765
step: 390, loss: 0.010382712818682194
step: 400, loss: 0.0017636247212067246
step: 410, loss: 0.0010548767168074846
step: 420, loss: 0.00041360649629496604
step: 430, loss: 0.04140536114573479
step: 440, loss: 0.05779218301177025
step: 450, loss: 0.0041770534589886665
step: 460, loss: 0.0028397911228239536
step: 470, loss: 0.08831615000963211
step: 480, loss: 0.013515996746718884
step: 490, loss: 0.01483701542019844
step: 500, loss: 0.002385334810242057
step: 510, loss: 0.02942504733800888
step: 520, loss: 0.001540657365694642
step: 530, loss: 0.0027348361909389496
/usr/local/lib/python3.7/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=0.9385474860335196, f1=0.9458955223880597,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014764509396627545
step: 10, loss: 0.0029944584239274263
step: 20, loss: 0.0012709477450698614
step: 30, loss: 0.10790468752384186
step: 40, loss: 0.0004460390191525221
step: 50, loss: 0.0006721894606016576
step: 60, loss: 0.001031251740641892
step: 70, loss: 0.005793807562440634
step: 80, loss: 0.016912769526243121
step: 90, loss: 0.08060289919376373
step: 100, loss: 0.009580142796039581
step: 110, loss: 0.012833629734814167
step: 120, loss: 0.020389854907989502
step: 130, loss: 0.0005269237444736063
step: 140, loss: 0.005811036564409733
step: 150, loss: 0.0004649549082387239
step: 160, loss: 0.06543136388063431
step: 170, loss: 0.010175053030252457
```

```
step: 180, loss: 0.00454481178894639
step: 190, loss: 0.004091838374733925
step: 200, loss: 0.0011446980061009526
step: 210, loss: 0.004332704935222864
step: 220, loss: 0.08087471127510071
step: 230, loss: 0.033698953688144684
step: 240, loss: 0.0005537302931770682
step: 250, loss: 0.0009718802757561207
step: 260, loss: 0.0003733987978193909
step: 270, loss: 0.020354371517896652
step: 280, loss: 0.003836799645796418
step: 290, loss: 0.00038180421688593924
step: 300, loss: 0.0005298759788274765
step: 310, loss: 0.158136248588562
step: 320, loss: 0.0030727828852832317
step: 330, loss: 0.004668518900871277
step: 340, loss: 0.012267941609025002
step: 350, loss: 0.06724395602941513
step: 360, loss: 0.0033210739493370056
step: 370, loss: 0.0007627354352734983
step: 380, loss: 0.0032236718107014894
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 390, loss: 0.00042659699101932347
step: 400, loss: 0.09810440242290497
step: 410, loss: 0.008809110149741173
step: 420, loss: 0.0013698602560907602
step: 430, loss: 0.05774417147040367
step: 440, loss: 0.000523307709954679
step: 450, loss: 0.0017607551999390125
step: 460, loss: 0.00014897863729856908
step: 470, loss: 0.0006701069651171565
step: 480, loss: 0.0001966050040209666
step: 490, loss: 0.12341208755970001
step: 500, loss: 0.0007874747971072793
step: 510, loss: 0.01689324714243412
step: 520, loss: 0.01284931879490614
step: 530, loss: 0.018281247466802597
/usr/local/lib/python3.7/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=0.9448373408769447, f1=0.943502824858757,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040343220462091267
step: 10, loss: 0.005346579011529684
step: 20, loss: 3.648400161182508e-05
```

```
step: 30, loss: 0.040633074939250946
step: 40, loss: 0.001840913319028914
step: 50, loss: 0.00013465546362567693
step: 60, loss: 0.000516955042257905
step: 70, loss: 0.00045115043758414686
step: 80, loss: 0.00771803455427289
step: 90, loss: 0.00040057554724626243
step: 100, loss: 0.0011076917871832848
step: 110, loss: 0.011793898418545723
step: 120, loss: 0.00010707107139751315
step: 130, loss: 0.0020066171418875456
step: 140, loss: 0.00011911132605746388
step: 150, loss: 0.024718042463064194
step: 160, loss: 0.028494907543063164
step: 170, loss: 0.007528518792241812
step: 180, loss: 0.056943848729133606
step: 190, loss: 0.00023749639512971044
step: 200, loss: 0.009909254498779774
step: 210, loss: 0.0004824964271392673
step: 220, loss: 0.00013951399887446314
step: 230, loss: 0.00011922753765247762
step: 240, loss: 0.0012773709604516625
step: 250, loss: 0.00028256146470084786
step: 260, loss: 0.00963936373591423
step: 270, loss: 0.0006412355578504503
step: 280, loss: 0.01201288215816021
step: 290, loss: 0.0019385083578526974
step: 300, loss: 0.010208390653133392
step: 310, loss: 0.050492893904447556
step: 320, loss: 0.019671112298965454
step: 330, loss: 0.02900168113410473
step: 340, loss: 0.010512545704841614
step: 350, loss: 0.0006638321792706847
step: 360, loss: 0.00019378263095859438
step: 370, loss: 0.006107087712734938
step: 380, loss: 0.0016721782740205526
step: 390, loss: 7.775568519718945e-05
step: 400, loss: 0.019772034138441086
step: 410, loss: 0.014602878130972385
step: 420, loss: 0.003471852047368884
step: 430, loss: 0.00012615995365194976
step: 440, loss: 0.0002250498509965837
step: 450, loss: 0.0013295251410454512
step: 460, loss: 0.0006577090243808925
step: 470, loss: 0.07819725573062897
step: 480, loss: 6.842787843197584e-05
step: 490, loss: 0.0029329576063901186
step: 500, loss: 0.0010996242053806782
```

```
step: 510, loss: 0.0024627423845231533
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 520, loss: 0.0036857326049357653
step: 530, loss: 0.0036357238423079252
/usr/local/lib/python3.7/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=0.9433255269320844, f1=0.9437675726335519,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013328625936992466
step: 10, loss: 0.05583162233233452
step: 20, loss: 0.0006781912525184453
step: 30, loss: 0.0020540368277579546
step: 40, loss: 0.0005141739966347814
step: 50, loss: 6.216007022885606e-05
step: 60, loss: 0.001206273678690195
step: 70, loss: 8.321466157212853e-05
step: 80, loss: 0.000510361201129854
step: 90, loss: 0.002755052875727415
step: 100, loss: 0.0010034845909103751
step: 110, loss: 5.1890299801016226e-05
step: 120, loss: 0.08308438956737518
step: 130, loss: 0.04260674864053726
step: 140, loss: 0.0006033065728843212
step: 150, loss: 0.0004992581089027226
step: 160, loss: 0.00313020427711308
step: 170, loss: 0.0038239655550569296
step: 180, loss: 0.0005879153031855822
step: 190, loss: 0.001483630621805787
step: 200, loss: 0.0002848173025995493
step: 210, loss: 0.0011003537802025676
step: 220, loss: 0.0020107279997318983
step: 230, loss: 0.000669653236400336
step: 240, loss: 0.0006889504729770124
step: 250, loss: 0.0015999863389879465
step: 260, loss: 0.0018124659545719624
step: 270, loss: 0.0006571935373358428
step: 280, loss: 0.0029708202928304672
step: 290, loss: 0.00999978557229042
step: 300, loss: 0.013610351830720901
step: 310, loss: 0.06438552588224411
step: 320, loss: 0.003919179085642099
step: 330, loss: 0.00011168696690001525
step: 340, loss: 0.059532467275857925
step: 350, loss: 0.0005217381403781474
```

```
step: 360, loss: 0.00034535658778622746
step: 370, loss: 0.0015646573156118393
step: 380, loss: 0.0022340859286487103
step: 390, loss: 0.0014083642745390534
step: 400, loss: 0.0001607566955499351
step: 410, loss: 0.000596068159211427
step: 420, loss: 0.00437001000344753
step: 430, loss: 0.002126468578353524
step: 440, loss: 0.0006508514052256942
step: 450, loss: 0.003182164393365383
step: 460, loss: 0.02531747706234455
step: 470, loss: 0.00012109026283724234
step: 480, loss: 0.00034465923090465367
step: 490, loss: 0.09044736623764038
step: 500, loss: 0.002313286531716585
step: 510, loss: 0.00034387651248835027
step: 520, loss: 0.005240664351731539
step: 530, loss: 0.007893714122474194
/usr/local/lib/python3.7/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=0.940677966101695, f1=0.948356807511737,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002182076219469309
step: 10, loss: 0.0037013692781329155
step: 20, loss: 0.00423989724367857
step: 30, loss: 0.11391467601060867
step: 40, loss: 0.005252205766737461
step: 50, loss: 0.00029635746614076197
step: 60, loss: 0.0006394809461198747
step: 70, loss: 0.0004593915946315974
step: 80, loss: 0.006456300150603056
step: 90, loss: 0.001422931207343936
step: 100, loss: 0.3054872453212738
step: 110, loss: 0.05212792009115219
step: 120, loss: 0.00010404278145870194
step: 130, loss: 0.0003926072677131742
step: 140, loss: 9.445215982850641e-05
step: 150, loss: 7.676560198888183e-05
step: 160, loss: 0.0022362309973686934
step: 170, loss: 0.001480895676650107
step: 180, loss: 0.011136211454868317
step: 190, loss: 0.0005657834117300808
step: 200, loss: 0.025949260219931602
step: 210, loss: 0.00027347682043910027
```

```
step: 220, loss: 0.0015680596698075533
step: 230, loss: 0.002042248146608472
step: 240, loss: 0.0014624171890318394
step: 250, loss: 6.106866203481331e-05
step: 260, loss: 5.202427564654499e-05
step: 270, loss: 0.001639775582589209
step: 280, loss: 0.0002665169013198465
step: 290, loss: 0.0021065841428935528
step: 300, loss: 0.001103721559047699
step: 310, loss: 0.0010111581068485975
step: 320, loss: 0.11117690056562424
step: 330, loss: 0.0013354275142773986
step: 340, loss: 0.12600058317184448
step: 350, loss: 0.0001359654706902802
step: 360, loss: 0.0003131514531560242
step: 370, loss: 0.0007259984849952161
step: 380, loss: 0.002612554468214512
step: 390, loss: 0.00011755896412068978
step: 400, loss: 4.5246943045640364e-05
step: 410, loss: 0.00031248515006154776
step: 420, loss: 0.005655864719301462
step: 430, loss: 0.0005915041547268629
step: 440, loss: 0.00033821642864495516
step: 450, loss: 0.0010385889327153563
step: 460, loss: 0.0003040010342374444
step: 470, loss: 0.032561782747507095
step: 480, loss: 0.07446634024381638
step: 490, loss: 0.0029491104651242495
step: 500, loss: 0.000237040949286893
step: 510, loss: 0.008562824688851833
step: 520, loss: 0.002782338298857212
step: 530, loss: 0.0003750088217202574
/usr/local/lib/python3.7/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=0.9420423183072677, f1=0.943758573388203,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014590880600735545
step: 10, loss: 0.0005575665272772312
step: 20, loss: 0.0004731877415906638
step: 30, loss: 0.00019188184523954988
step: 40, loss: 0.10176370292901993
step: 50, loss: 0.003034807275980711
step: 60, loss: 0.0006160925840958953
step: 70, loss: 0.0025882218033075333
```

```
step: 80, loss: 0.009047139436006546
step: 90, loss: 0.00019816015264950693
step: 100, loss: 3.956177897634916e-05
step: 110, loss: 2.5010946046677418e-05
step: 120, loss: 0.0015034672105684876
step: 130, loss: 7.590877066832036e-05
step: 140, loss: 0.21019162237644196
step: 150, loss: 0.0022742371074855328
step: 160, loss: 0.002108306623995304
step: 170, loss: 0.0013575643533840775
step: 180, loss: 0.0004307953058741987
step: 190, loss: 0.001937403460033238
step: 200, loss: 0.20699165761470795
step: 210, loss: 0.0003142217465210706
step: 220, loss: 0.00034969483385793865
step: 230, loss: 0.002156591974198818
step: 240, loss: 0.0007222738349810243
step: 250, loss: 0.0003446001501288265
step: 260, loss: 7.968347199494019e-05
step: 270, loss: 0.002162154531106353
step: 280, loss: 0.0002628003421705216
step: 290, loss: 0.0028814903926104307
step: 300, loss: 7.738162821624428e-05
step: 310, loss: 0.0007252370123751462
step: 320, loss: 0.00031153816962614655
step: 330, loss: 0.017303019762039185
step: 340, loss: 0.010732880793511868
step: 350, loss: 6.284491973929107e-05
step: 360, loss: 0.03242804855108261
step: 370, loss: 0.0025508373510092497
step: 380, loss: 0.0001861509372247383
step: 390, loss: 0.00014797825133427978
step: 400, loss: 0.0001246391038876027
step: 410, loss: 7.614852074766532e-05
step: 420, loss: 0.00011528359027579427
step: 430, loss: 0.0001321369200013578
step: 440, loss: 0.003009201493114233
step: 450, loss: 0.0017757996683940291
step: 460, loss: 0.023810993880033493
step: 470, loss: 0.0029372444842010736
step: 480, loss: 1.814559436752461e-05
step: 490, loss: 0.0006997293094173074
step: 500, loss: 0.00019743139273487031
step: 510, loss: 0.0006058515864424407
step: 520, loss: 0.0007600327371619642
step: 530, loss: 0.00015277741476893425
/usr/local/lib/python3.7/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=0.9425815342214057, f1=0.9431714023831348,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023097859229892492
step: 10, loss: 0.00026754080317914486
step: 20, loss: 5.386997509049252e-05
step: 30, loss: 0.0015624144580215216
step: 40, loss: 0.0005341527285054326
step: 50, loss: 0.0031951339915394783
step: 60, loss: 0.006396890617907047
step: 70, loss: 0.005423145368695259
step: 80, loss: 0.0006073086406104267
step: 90, loss: 5.818307545268908e-05
step: 100, loss: 0.003103457624092698
step: 110, loss: 0.0019753617234528065
step: 120, loss: 0.00016007859085220844
step: 130, loss: 0.0003873104287777096
step: 140, loss: 0.0024233381263911724
step: 150, loss: 0.010676147416234016
step: 160, loss: 0.0007655789377167821
step: 170, loss: 0.0009976584697142243
step: 180, loss: 0.00016525365936104208
step: 190, loss: 8.116529352264479e-05
step: 200, loss: 4.126426574657671e-05
step: 210, loss: 0.0007495511672459543
step: 220, loss: 0.00011099781113443896
step: 230, loss: 0.008714464493095875
step: 240, loss: 0.00013494383892975748
step: 250, loss: 0.0014998277183622122
step: 260, loss: 0.001025347737595439
step: 270, loss: 0.00010375284182373434
step: 280, loss: 6.777640373911709e-05
step: 290, loss: 9.330164903076366e-05
step: 300, loss: 0.00016364653129130602
step: 310, loss: 0.0007211302872747183
step: 320, loss: 6.367718742694706e-05
step: 330, loss: 0.00013852295523975044
step: 340, loss: 0.00028297241078689694
step: 350, loss: 0.00012765222345478833
step: 360, loss: 3.6025387089466676e-05
step: 370, loss: 0.005270163994282484
step: 380, loss: 0.0002784098032861948
step: 390, loss: 0.001150673720985651
step: 400, loss: 0.0029415299650281668
step: 410, loss: 2.27868677029619e-05
```

```
step: 420, loss: 0.03064429573714733
step: 430, loss: 0.0024337503127753735
step: 440, loss: 0.0007028175750747323
step: 450, loss: 1.9769569917116314e-05
step: 460, loss: 0.02539634518325329
step: 470, loss: 2.1814819774590433e-05
step: 480, loss: 0.00778642063960433
step: 490, loss: 0.002959928009659052
step: 500, loss: 0.00842248648405075
step: 510, loss: 0.006577291991561651
step: 520, loss: 0.00048345859977416694
step: 530, loss: 0.000285898830043152
/usr/local/lib/python3.7/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=0.9441860465116279, f1=0.9452181987000929,
best_f1=0.9406858202038925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031293820939026773
step: 10, loss: 0.0006090229144319892
step: 20, loss: 0.0004064243985339999
step: 30, loss: 0.015840308740735054
step: 40, loss: 0.00012957013677805662
step: 50, loss: 0.0005809771828353405
step: 60, loss: 0.016020653769373894
step: 70, loss: 8.492809138260782e-05
step: 80, loss: 0.0001906873076222837
step: 90, loss: 5.1220642490079626e-05
step: 100, loss: 1.89948441402521e-05
step: 110, loss: 0.00017606366600375623
step: 120, loss: 5.336132744560018e-05
step: 130, loss: 0.0014033436309546232
step: 140, loss: 5.238685480435379e-05
step: 150, loss: 0.0018268655985593796
step: 160, loss: 1.5463443560292944e-05
step: 170, loss: 6.556487642228603e-05
step: 180, loss: 0.000321311061270535
step: 190, loss: 0.0003667631244752556
step: 200, loss: 0.00022814182739239186
step: 210, loss: 0.0005450028111226857
step: 220, loss: 0.00030002923449501395
step: 230, loss: 0.001461623003706336
step: 240, loss: 0.0015668139094486833
step: 250, loss: 0.00010426133667351678
step: 260, loss: 1.319111015618546e-05
step: 270, loss: 3.776044468395412e-05
```

```
step: 280, loss: 0.0006939050508663058
step: 290, loss: 0.0004704951134044677
step: 300, loss: 0.0008475988870486617
step: 310, loss: 0.0012141010956838727
step: 320, loss: 0.0002001743414439261
step: 330, loss: 5.096737731946632e-05
step: 340, loss: 3.105957875959575e-05
step: 350, loss: 0.003985963761806488
step: 360, loss: 0.0001886010286398232
step: 370, loss: 0.0033930393401533365
step: 380, loss: 0.0016071011777967215
step: 390, loss: 1.5914223695290275e-05
step: 400, loss: 0.07831664383411407
step: 410, loss: 3.2351366826333106e-05
step: 420, loss: 8.546026219846681e-05
step: 430, loss: 1.843982499849517e-05
step: 440, loss: 0.005094294436275959
step: 450, loss: 0.0025658253580331802
step: 460, loss: 0.00041990194586105645
step: 470, loss: 0.002449811901897192
step: 480, loss: 0.00013963245146442205
step: 490, loss: 0.001109360484406352
step: 500, loss: 0.0004843553469981998
step: 510, loss: 3.968795863329433e-05
step: 520, loss: 1.6335163309122436e-05
step: 530, loss: 0.0004989312728866935
/usr/local/lib/python3.7/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=0.9429906542056075, f1=0.9468283582089553,
best_f1=0.9406858202038925
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 [00:34, 167.09it/s]
load_f1 = 0.9480459770114943
real_f1 = 0.9485766758494032
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 142.45it/s]
```