

ADirty_90_3_5_bert

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 23.72 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.4 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 56.5 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 53.4 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.7 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 66.0 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 31.3 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 22.70 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
|                                     | 10.4 MB 3.6 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 52.2 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 2.4 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 28.0 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 57.9 MB/s
Requirement already satisfied: scipy>=0.18.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.7.3)
Requirement already satisfied: 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: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    |                               | 184 kB 70.9 MB/s
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    |                               | 2.2 MB 50.0 MB/s
Collecting blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
    |                               | 3.7 MB 48.9 MB/s
Requirement already satisfied: setuptools in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (57.4.0)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)

```

```

| 880 kB 53.8 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    | 3.0 MB 56.6 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: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!<1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449923
sha256=9e2251f61747fa5b7535a92c60a3d3dd2f1853ee35cdb5502e28a99d2b5f8c46
  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=8017edeac256484cdfc23c98624971d72a9a2b92ae8badcf0271cbb11dbfb505
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 118), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.96 MiB | 27.51 MiB/s, done.
Resolving deltas: 100% (6922/6922), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writable
Created temporary directory: /tmp/pip-ephem-wheel-cache-5wffcd1b
Created temporary directory: /tmp/pip-req-tracker-8hkpjas0
Initialized build tracking at /tmp/pip-req-tracker-8hkpjas0
Created build tracker: /tmp/pip-req-tracker-8hkpjas0
Entered build tracker: /tmp/pip-req-tracker-8hkpjas0
```

```

Created temporary directory: /tmp/pip-install-spd1krn_
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-_hs43qmk
  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-8hkpjas0'
  Running setup.py (path:/tmp/pip-req-build-_hs43qmk/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-205g9gqt
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-205g9gqt/apex.egg-info
  writing /tmp/pip-pip-egg-info-205g9gqt/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-205g9gqt/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-205g9gqt/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-205g9gqt/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-205g9gqt/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-_hs43qmk 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-8hkpjas0'
Created temporary directory: /tmp/pip-unpack-ed66ry_
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-jv3tllm0
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-jv3tllm0
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-_hs43qmk/setup.py''';

```



```
__file__='''/tmp/pip-req-build-_hs43qmk/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-jv3tllm0
```

```
torch.__version__ = 1.4.0
```

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

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```



```

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

```

```

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

```

```

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

```

```

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

```

```

    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    running install_egg_info
    running egg_info
    creating apex.egg-info
    writing apex.egg-info/PKG-INFO
    writing dependency_links to apex.egg-info/dependency_links.txt
    writing top-level names to apex.egg-info/top_level.txt
    writing manifest file 'apex.egg-info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file 'apex.egg-info/SOURCES.txt'
    Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
    running install_scripts
    adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
    creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
    creating '/tmp/pip-wheel-jv3tllm0/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
    adding 'apex/__init__.py'
    adding 'apex/_autocast_utils.py'
    adding 'apex/RNN/RNNBackend.py'

```

```

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

```

```

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

```

```

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

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298244
sha256=8d0fcea2835ce2f8e38770d0fa4113ddf166fff96c1b2a281ffeac9e656be717
Stored in directory: /tmp/pip-ephem-wheel-cache-5wffcd1b/wheels/6b/22/a2/16655
26ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-8hkpjas0'
/content/ditto

```

```

[3]: # 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 3.2 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 24.3 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 10.5 MB/s
Collecting botocore==1.27.48
  Downloading botocore-1.27.48-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 47.8 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 colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 36.3 MB/s
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 17 kB/s
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages
(from torch==1.7.1) (1.19.2)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0
```

```

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

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

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

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

```

0.3 importa config.json

```
[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 | 31.04 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/datasasErros/ADirty_90_3_5/configs.json /content/ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

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

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

Downloading: 100% 433/433 [00:00<00:00, 477kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 21.7MB/s]

Downloading: 100% 440M/440M [00:10<00:00, 41.5MB/s]

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

Defaults for this optimization level are:

enabled : True

opt_level : 02

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled               : True
opt_level             : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5720033049583435
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.34146341463414637, f1=0.32558139534883723,
best_f1=0.32558139534883723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5069850087165833
/usr/local/lib/python3.7/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=0.26666666666666667, f1=0.26666666666666667,

```

```

best_f1=0.32558139534883723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5798655152320862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5, f1=0.4761904761904762, best_f1=0.4761904761904762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3368094861507416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5384615384615384, f1=0.39285714285714285,
best_f1=0.39285714285714285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3068941533565521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.689655172413793, 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: 0.24506069719791412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.6666666666666666, f1=0.588235294117647,
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: 0.30417242646217346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7741935483870968, f1=0.6285714285714286,
best_f1=0.6285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09796712547540665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.6470588235294117,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0074877869337797165
/usr/local/lib/python3.7/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=0.6666666666666667, f1=0.6111111111111112,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07175976037979126
/usr/local/lib/python3.7/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=0.6086956521739131, f1=0.6086956521739131,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024537667632102966

```



```

/usr/local/lib/python3.7/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=0.6875000000000001, f1=0.689655172413793,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012160641141235828
/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.689655172413793,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03215481713414192
/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.6451612903225806,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005913224071264267
/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.6451612903225806,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021892879158258438
/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.6451612903225806,
best_f1=0.6470588235294117
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 137394.41it/s]
/usr/local/lib/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 = 0.7222222222222223
real_f1 = 0.7222222222222223
0it [00:00, ?it/s]/usr/local/lib/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 [00:14, 297.56it/s]

```

1.1.2 DBLP-ACM - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

```
step: 0, loss: 0.5949347019195557
step: 10, loss: 0.6474952101707458
step: 20, loss: 0.47262078523635864
step: 30, loss: 0.2920142412185669
step: 40, loss: 0.24742825329303741
step: 50, loss: 0.18259882926940918
step: 60, loss: 0.03269226849079132
step: 70, loss: 0.03941121697425842
step: 80, loss: 0.12630808353424072
step: 90, loss: 0.13624577224254608
step: 100, loss: 0.00811282079666853
step: 110, loss: 0.255650132894516
step: 120, loss: 0.010322793386876583
step: 130, loss: 0.02224976383149624
step: 140, loss: 0.002848349278792739
step: 150, loss: 0.05224922299385071
step: 160, loss: 0.03262464329600334
step: 170, loss: 0.03923500329256058
```

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

```
step: 180, loss: 0.027710745111107826
step: 190, loss: 0.10529842972755432
step: 200, loss: 0.12410972267389297
step: 210, loss: 0.0037677818909287453
step: 220, loss: 0.023919792845845222
step: 230, loss: 0.08862630277872086
```

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

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

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

```
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=0.9731543624161074, f1=0.9694915254237287,
best_f1=0.9694915254237287
```

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

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

```
step: 0, loss: 0.005354568362236023
step: 10, loss: 0.003106033429503441
step: 20, loss: 0.0682389885187149
step: 30, loss: 0.19083403050899506
```

```

step: 40, loss: 0.04039183631539345
step: 50, loss: 0.00621689623221755
step: 60, loss: 0.00247971061617136
step: 70, loss: 0.061013251543045044
step: 80, loss: 0.0073066093027591705
step: 90, loss: 0.003449326381087303
step: 100, loss: 0.0019724618177860975
step: 110, loss: 0.02771984227001667
step: 120, loss: 0.0011020053643733263
step: 130, loss: 0.008377890102565289
step: 140, loss: 0.052343934774398804
step: 150, loss: 0.006693798117339611
step: 160, loss: 0.003414308186620474
step: 170, loss: 0.029470132663846016
step: 180, loss: 0.014826801605522633
step: 190, loss: 0.034943535923957825
step: 200, loss: 0.002897003898397088
step: 210, loss: 0.0014401667285710573
step: 220, loss: 0.04498720541596413
step: 230, loss: 0.05995551869273186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9761634506242906,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016241326928138733
step: 10, loss: 0.003695590188726783
step: 20, loss: 0.004647485911846161
step: 30, loss: 0.010324399918317795
step: 40, loss: 0.12134087830781937
step: 50, loss: 0.05007064342498779
step: 60, loss: 0.0016660374822095037
step: 70, loss: 0.007403544150292873
step: 80, loss: 0.00044265229371376336
step: 90, loss: 0.18829013407230377
step: 100, loss: 0.012795482762157917
step: 110, loss: 0.0014464114792644978
step: 120, loss: 0.04524993151426315
step: 130, loss: 0.02004918083548546
step: 140, loss: 0.11190776526927948
step: 150, loss: 0.00989333726465702

```

```

step: 160, loss: 0.007005855906754732
step: 170, loss: 0.005971415434032679
step: 180, loss: 0.006063604727387428
step: 190, loss: 0.021255914121866226
step: 200, loss: 0.006006855517625809
step: 210, loss: 0.0010170917958021164
step: 220, loss: 0.0004989688168279827
step: 230, loss: 0.00031097972532734275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.983050847457627, f1=0.979591836734694,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038714613765478134
step: 10, loss: 0.00041800676262937486
step: 20, loss: 0.0011831915471702814
step: 30, loss: 0.00018742481188382953
step: 40, loss: 0.0011302692582830787
step: 50, loss: 0.0012500013690441847
step: 60, loss: 0.03746064007282257
step: 70, loss: 0.0013669467298313975
step: 80, loss: 0.006473298650234938
step: 90, loss: 0.001825847546570003
step: 100, loss: 0.0010409692768007517
step: 110, loss: 0.0002771923318505287
step: 120, loss: 0.005929472390562296
step: 130, loss: 0.001017891801893711
step: 140, loss: 0.0024423429276794195
step: 150, loss: 0.19095559418201447
step: 160, loss: 0.024743184447288513
step: 170, loss: 0.014330173842608929
step: 180, loss: 0.0009175604791380465
step: 190, loss: 0.0021322693210095167
step: 200, loss: 0.0012822240823879838
step: 210, loss: 0.01689569093286991
step: 220, loss: 0.0004453261790331453
step: 230, loss: 0.002721893833950162
/usr/local/lib/python3.7/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=0.9748858447488584, f1=0.9725400457665903,

```

```

best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007229694747366011
step: 10, loss: 0.0005583651945926249
step: 20, loss: 0.006039253436028957
step: 30, loss: 0.00152622873429209
step: 40, loss: 0.001277233357541263
step: 50, loss: 0.0003048629150725901
step: 60, loss: 0.0016674257349222898
step: 70, loss: 0.0007267402834258974
step: 80, loss: 0.02390642650425434
step: 90, loss: 0.0009264365071430802
step: 100, loss: 0.00044907984556630254
step: 110, loss: 0.00017362575454171747
step: 120, loss: 0.00010432703129481524
step: 130, loss: 0.0011732025304809213
step: 140, loss: 0.005150360055267811
step: 150, loss: 0.011823872104287148
step: 160, loss: 0.04138091951608658
step: 170, loss: 0.014091567136347294
step: 180, loss: 0.1237744614481926
step: 190, loss: 0.0018536016577854753
step: 200, loss: 0.002846787218004465
step: 210, loss: 0.003951105289161205
step: 220, loss: 0.01734408363699913
step: 230, loss: 0.001153103425167501
/usr/local/lib/python3.7/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=0.9753363228699552, f1=0.972972972972973,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013344367034733295
step: 10, loss: 0.0010986811248585582
step: 20, loss: 0.0067760637030005455
step: 30, loss: 0.0015524406917393208
step: 40, loss: 0.030252806842327118
step: 50, loss: 0.005761909764260054
step: 60, loss: 0.0014886637218296528
step: 70, loss: 0.0003865601902361959
step: 80, loss: 0.0012946303468197584
step: 90, loss: 0.003052087966352701
step: 100, loss: 0.1077704131603241
step: 110, loss: 0.0013454598374664783

```

```

step: 120, loss: 0.0031285129953175783
step: 130, loss: 0.009500396437942982
step: 140, loss: 0.001330572529695928
step: 150, loss: 0.0017119050025939941
step: 160, loss: 0.002214224310591817
step: 170, loss: 0.0007530316361226141
step: 180, loss: 0.03469209000468254
step: 190, loss: 0.04997289553284645
step: 200, loss: 0.0043792566284537315
step: 210, loss: 0.006627856288105249
step: 220, loss: 0.021487068384885788
step: 230, loss: 0.09391648322343826
/usr/local/lib/python3.7/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=0.9709821428571428, f1=0.972972972972973,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007179269101470709
step: 10, loss: 0.00022568897111341357
step: 20, loss: 0.0014499187236651778
step: 30, loss: 0.0004356866411399096
step: 40, loss: 0.001181766507215798
step: 50, loss: 0.0004927815753035247
step: 60, loss: 0.024932876229286194
step: 70, loss: 0.01567830517888069
step: 80, loss: 0.0012657769257202744
step: 90, loss: 0.00016317566041834652
step: 100, loss: 0.00025059605832211673
step: 110, loss: 0.0006649639108218253
step: 120, loss: 0.0002695767907425761
step: 130, loss: 0.00026473807520233095
step: 140, loss: 0.00044195743976160884
step: 150, loss: 0.0003579641634132713
step: 160, loss: 0.1112314835190773
step: 170, loss: 0.015542935580015182
step: 180, loss: 0.0009067466016858816
step: 190, loss: 0.00097687728703022
step: 200, loss: 0.027908772230148315
step: 210, loss: 0.00016536891052965075
step: 220, loss: 0.003762422129511833
step: 230, loss: 0.0002570563810877502
/usr/local/lib/python3.7/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=0.9744160177975528, f1=0.9655937846836848,

```



```

best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001488400885136798
step: 10, loss: 0.00040165966493077576
step: 20, loss: 0.00010290925274603069
step: 30, loss: 0.00015587943198624998
step: 40, loss: 0.0008212769753299654
step: 50, loss: 0.00498042069375515
step: 60, loss: 0.00032505931449122727
step: 70, loss: 0.0007154599297791719
step: 80, loss: 0.004198855720460415
step: 90, loss: 0.00010747288615675643
step: 100, loss: 0.0005486739100888371
step: 110, loss: 0.00039421196561306715
step: 120, loss: 0.00024886461324058473
step: 130, loss: 0.04313312843441963
step: 140, loss: 0.00019223896379116923
step: 150, loss: 0.0005424689734354615
step: 160, loss: 0.00016021894407458603
step: 170, loss: 0.00017904816195368767
step: 180, loss: 0.00028651917818933725
step: 190, loss: 0.0001977930951397866
step: 200, loss: 0.0037150857970118523
step: 210, loss: 0.00025709872716106474
step: 220, loss: 0.001629595528356731
step: 230, loss: 0.0001702653826214373
/usr/local/lib/python3.7/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=0.9775784753363228, f1=0.9731543624161074,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.78300886345096e-05
step: 10, loss: 0.0003629827988334
step: 20, loss: 0.0006425310275517404
step: 30, loss: 9.526211943011731e-05
step: 40, loss: 0.037043143063783646
step: 50, loss: 0.00030971551313996315
step: 60, loss: 0.00027338560903444886
step: 70, loss: 0.00017517055675853044
step: 80, loss: 0.0010211883345618844
step: 90, loss: 0.0001617055240785703
step: 100, loss: 0.0073409490287303925
step: 110, loss: 8.042235276661813e-05

```

```

step: 120, loss: 6.363471038639545e-05
step: 130, loss: 9.473100362811238e-05
step: 140, loss: 9.506390051683411e-05
step: 150, loss: 0.00044064680696465075
step: 160, loss: 7.906159589765593e-05
step: 170, loss: 0.00019091971626039594
step: 180, loss: 0.0012545585632324219
step: 190, loss: 0.00015008858463261276
step: 200, loss: 0.00018672032456379384
step: 210, loss: 0.0012927788775414228
step: 220, loss: 0.000591087038628757
step: 230, loss: 0.0017139182891696692
/usr/local/lib/python3.7/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=0.9705215419501134, f1=0.9727272727272728,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005258644232526422
step: 10, loss: 5.5477008572779596e-05
step: 20, loss: 0.00036613992415368557
step: 30, loss: 0.00019674631766974926
step: 40, loss: 0.00010898730397457257
step: 50, loss: 0.00013211106124799699
step: 60, loss: 0.000933364499360323
step: 70, loss: 0.0005846096319146454
step: 80, loss: 0.00018347262812312692
step: 90, loss: 0.00040247265133075416
step: 100, loss: 0.00011314541916362941
step: 110, loss: 0.00013017418677918613
step: 120, loss: 0.0010237437672913074
step: 130, loss: 6.929630762897432e-05
step: 140, loss: 0.053057827055454254
step: 150, loss: 0.004924169275909662
step: 160, loss: 4.4042975787306204e-05
step: 170, loss: 8.740316843613982e-05
step: 180, loss: 0.00025300614652223885
step: 190, loss: 0.022930286824703217
step: 200, loss: 7.266853208420798e-05
step: 210, loss: 7.417523011099547e-05
step: 220, loss: 0.0002572553639765829
step: 230, loss: 0.00012190519919386134
/usr/local/lib/python3.7/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=0.9766925638179801, f1=0.9787234042553192,

```

```

best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027129726368002594
step: 10, loss: 7.247338362503797e-05
step: 20, loss: 0.0032187344040721655
step: 30, loss: 0.02081144042313099
step: 40, loss: 0.0001835258590290323
step: 50, loss: 5.1741029892582446e-05
step: 60, loss: 0.006765812169760466
step: 70, loss: 7.11529646650888e-05
step: 80, loss: 6.103933264967054e-05
step: 90, loss: 9.855032840278e-05
step: 100, loss: 6.309230229817331e-05
step: 110, loss: 6.532757834065706e-05
step: 120, loss: 8.323500514961779e-05
step: 130, loss: 7.443549111485481e-05
step: 140, loss: 0.0005015701754018664
step: 150, loss: 0.058118946850299835
step: 160, loss: 5.690529360435903e-05
step: 170, loss: 0.0014442243846133351
step: 180, loss: 0.00014189672947395593
step: 190, loss: 5.183899702387862e-05
step: 200, loss: 6.651139847235754e-05
step: 210, loss: 5.377221168600954e-05
step: 220, loss: 8.497143426211551e-05
step: 230, loss: 9.284912084694952e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9797297297297298, f1=0.9763779527559054,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.028716936474666e-05
step: 10, loss: 7.577845826745033e-05
step: 20, loss: 0.00014419904618989676
step: 30, loss: 0.0038363554049283266
step: 40, loss: 7.433644350385293e-05
step: 50, loss: 0.006244224030524492
step: 60, loss: 0.0010457466123625636
step: 70, loss: 5.104025694890879e-05
step: 80, loss: 5.831171802128665e-05
step: 90, loss: 5.573937596636824e-05
step: 100, loss: 0.0004397466836962849
step: 110, loss: 8.027235890040174e-05

```

```

step: 120, loss: 0.00011214467667741701
step: 130, loss: 0.0001484107633586973
step: 140, loss: 9.388761827722192e-05
step: 150, loss: 0.00010268465848639607
step: 160, loss: 0.0001870602573035285
step: 170, loss: 8.131224603857845e-05
step: 180, loss: 9.530560055281967e-05
step: 190, loss: 0.0002210782840847969
step: 200, loss: 5.495607547345571e-05
step: 210, loss: 6.488856161013246e-05
step: 220, loss: 0.0001257963594980538
step: 230, loss: 0.00020813783339690417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831271091113611, f1=0.9786276715410572,
best_f1=0.9786276715410572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.314345177495852e-05
step: 10, loss: 7.501692743971944e-05
step: 20, loss: 0.00024913757806643844
step: 30, loss: 0.0002564121095929295
step: 40, loss: 0.00010590069723548368
step: 50, loss: 0.0005255905562080443
step: 60, loss: 0.00011051712499465793
step: 70, loss: 4.164519123150967e-05
step: 80, loss: 7.719955465290695e-05
step: 90, loss: 0.0003762313863262534
step: 100, loss: 2.7260606657364406e-05
step: 110, loss: 0.011450082994997501
step: 120, loss: 0.0007605403661727905
step: 130, loss: 9.741722169565037e-05
step: 140, loss: 5.929883263888769e-05
step: 150, loss: 5.380453512771055e-05
step: 160, loss: 0.005363497883081436
step: 170, loss: 8.756366150919348e-05
step: 180, loss: 6.672771996818483e-05
step: 190, loss: 4.916414036415517e-05
step: 200, loss: 0.00038068380672484636
step: 210, loss: 3.482116881059483e-05
step: 220, loss: 0.000483106094179675
step: 230, loss: 5.917581802350469e-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=0.9830124575311437, f1=0.9751131221719457,
best_f1=0.9786276715410572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.6116943596862257e-05
step: 10, loss: 0.006677971687167883
step: 20, loss: 4.971312591806054e-05
step: 30, loss: 8.037830411922187e-05
step: 40, loss: 0.0007441797060891986
step: 50, loss: 9.350258915219456e-05
step: 60, loss: 0.00014386010298039764
step: 70, loss: 0.00025639741215854883
step: 80, loss: 5.723380672861822e-05
step: 90, loss: 8.812348096398637e-05
step: 100, loss: 4.8284015065291896e-05
step: 110, loss: 0.00016986938135232776
step: 120, loss: 2.7357580620446242e-05
step: 130, loss: 5.902600969420746e-05
step: 140, loss: 7.611369801452383e-05
step: 150, loss: 4.938183701597154e-05
step: 160, loss: 0.008292661048471928
step: 170, loss: 2.3275202693184838e-05
step: 180, loss: 6.114703137427568e-05
step: 190, loss: 6.4228726841975e-05
step: 200, loss: 6.343552377074957e-05
step: 210, loss: 0.0002187188365496695
step: 220, loss: 7.336428097914904e-05
step: 230, loss: 6.618392944801599e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9830890642615557, f1=0.9741282339707535,
best_f1=0.9786276715410572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.6960725740063936e-05
step: 10, loss: 3.345114237163216e-05
step: 20, loss: 7.172395999077708e-05
step: 30, loss: 8.460569370072335e-05
step: 40, loss: 0.00012680955114774406
step: 50, loss: 0.04565681144595146
step: 60, loss: 5.245486318017356e-05
step: 70, loss: 0.0001290431246161461

```

```

step: 80, loss: 0.0001006711318041198
step: 90, loss: 8.393786265514791e-05
step: 100, loss: 0.00012114617129554972
step: 110, loss: 6.350325566018e-05
step: 120, loss: 7.22353724995628e-05
step: 130, loss: 6.498800939880311e-05
step: 140, loss: 4.927697227685712e-05
step: 150, loss: 0.00018738643848337233
step: 160, loss: 4.94109008286614e-05
step: 170, loss: 5.0510319852037355e-05
step: 180, loss: 0.0059735774993896484
step: 190, loss: 6.138319440651685e-05
step: 200, loss: 6.554508581757545e-05
step: 210, loss: 6.371762719936669e-05
step: 220, loss: 5.48948701180052e-05
step: 230, loss: 6.633638258790597e-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=0.983050847457627, f1=0.972972972972973,
best_f1=0.9786276715410572
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:11, 211.46it/s]
load_f1 = 0.9785310734463276
real_f1 = 0.9740112994350283
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:16, 269.68it/s]

```

1.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

```

Downloading: 100% 433/433 [00:00<00:00, 393kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 1.23MB/s]

```

Downloading: 100% 440M/440M [00:09<00:00, 48.6MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

step: 0, loss: 0.647811233997345
step: 10, loss: 0.5494239926338196
step: 20, loss: 0.5842128396034241
step: 30, loss: 0.36220595240592957
step: 40, loss: 0.1723824143409729
step: 50, loss: 0.28127768635749817
step: 60, loss: 0.09187918901443481
step: 70, loss: 0.11512069404125214
step: 80, loss: 0.04243184253573418
step: 90, loss: 0.5786586999893188
step: 100, loss: 0.14049172401428223


```
step: 110, loss: 0.11035636812448502
step: 120, loss: 0.1745450645685196
step: 130, loss: 0.09030871093273163
step: 140, loss: 0.10694514214992523
step: 150, loss: 0.11331816017627716
step: 160, loss: 0.051772214472293854
step: 170, loss: 0.25438928604125977
step: 180, loss: 0.1362900286912918
step: 190, loss: 0.06452261656522751
step: 200, loss: 0.1307314932346344
step: 210, loss: 0.09265559166669846
step: 220, loss: 0.26909446716308594
step: 230, loss: 0.16347548365592957
step: 240, loss: 0.08283238112926483
step: 250, loss: 0.1285710632801056
step: 260, loss: 0.3250022232532501
step: 270, loss: 0.031790196895599365
step: 280, loss: 0.100617416203022
step: 290, loss: 0.10284170508384705
step: 300, loss: 0.14861708879470825
step: 310, loss: 0.14149491488933563
step: 320, loss: 0.08952653408050537
step: 330, loss: 0.05216789245605469
step: 340, loss: 0.1644037365913391
step: 350, loss: 0.08751174807548523
step: 360, loss: 0.03602233901619911
step: 370, loss: 0.19372649490833282
step: 380, loss: 0.02706661820411682
step: 390, loss: 0.34010806679725647
step: 400, loss: 0.2150147557258606
step: 410, loss: 0.1426236480474472
step: 420, loss: 0.14524458348751068
step: 430, loss: 0.15586505830287933
step: 440, loss: 0.03584467992186546
step: 450, loss: 0.0261545330286026
step: 460, loss: 0.02530554123222828
step: 470, loss: 0.09829290956258774
step: 480, loss: 0.1003790944814682
step: 490, loss: 0.24213235080242157
step: 500, loss: 0.10277579724788666
step: 510, loss: 0.14047175645828247
step: 520, loss: 0.028476187959313393
step: 530, loss: 0.015119875781238079
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8903406439570696, f1=0.889814814814815,
best_f1=0.889814814814815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16397078335285187
step: 10, loss: 0.10100184381008148
step: 20, loss: 0.010834301821887493
step: 30, loss: 0.010852339677512646
step: 40, loss: 0.10155065357685089
step: 50, loss: 0.1616738736629486
step: 60, loss: 0.04857989773154259
step: 70, loss: 0.13431353867053986
step: 80, loss: 0.20593476295471191
step: 90, loss: 0.055593740195035934
step: 100, loss: 0.1266070455312729
step: 110, loss: 0.04535534977912903
step: 120, loss: 0.15490931272506714
step: 130, loss: 0.08612284064292908
step: 140, loss: 0.05870300531387329
step: 150, loss: 0.0986516997218132
step: 160, loss: 0.10018251836299896
step: 170, loss: 0.06850933283567429
step: 180, loss: 0.049243755638599396
step: 190, loss: 0.04854971170425415
step: 200, loss: 0.04228592664003372
step: 210, loss: 0.11171203851699829
step: 220, loss: 0.044274769723415375
step: 230, loss: 0.018724270164966583
step: 240, loss: 0.037176717072725296
step: 250, loss: 0.09846056997776031
step: 260, loss: 0.03338748589158058
step: 270, loss: 0.2726437449455261
step: 280, loss: 0.08380959928035736
step: 290, loss: 0.0708615854382515
step: 300, loss: 0.0588870607316494
step: 310, loss: 0.02816847711801529
step: 320, loss: 0.24990598857402802
step: 330, loss: 0.05557199567556381
step: 340, loss: 0.014593265950679779
step: 350, loss: 0.01424410566687584
step: 360, loss: 0.11928031593561172
step: 370, loss: 0.12069493532180786
step: 380, loss: 0.2223578691482544
step: 390, loss: 0.12777690589427948
step: 400, loss: 0.02957611158490181

```

```

step: 410, loss: 0.024338027462363243
step: 420, loss: 0.08180277049541473
step: 430, loss: 0.04348449781537056
step: 440, loss: 0.052223753184080124
step: 450, loss: 0.038351885974407196
step: 460, loss: 0.056531406939029694
step: 470, loss: 0.017820097506046295
step: 480, loss: 0.1492721289396286
step: 490, loss: 0.04218794032931328
step: 500, loss: 0.37766870856285095
step: 510, loss: 0.035944439470767975
step: 520, loss: 0.12610788643360138
step: 530, loss: 0.05543491244316101
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9072261072261073, f1=0.9066293183940244,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07293684035539627
step: 10, loss: 0.028563128784298897
step: 20, loss: 0.2432786375284195
step: 30, loss: 0.03760824352502823
step: 40, loss: 0.012613344937562943
step: 50, loss: 0.014415017329156399
step: 60, loss: 0.005697699263691902
step: 70, loss: 0.04613369703292847
step: 80, loss: 0.003329269355162978
step: 90, loss: 0.017527220770716667
step: 100, loss: 0.06298162043094635
step: 110, loss: 0.0048758178018033504
step: 120, loss: 0.051425911486148834
step: 130, loss: 0.005377823952585459
step: 140, loss: 0.027138348668813705
step: 150, loss: 0.06164112687110901
step: 160, loss: 0.15166126191616058
step: 170, loss: 0.1408804953098297
step: 180, loss: 0.01747901737689972
step: 190, loss: 0.059946998953819275
step: 200, loss: 0.04112037271261215
step: 210, loss: 0.02553158439695835
step: 220, loss: 0.17295090854167938

```

```

step: 230, loss: 0.07649321109056473
step: 240, loss: 0.030881347134709358
step: 250, loss: 0.13979646563529968
step: 260, loss: 0.03484697639942169
step: 270, loss: 0.011363185942173004
step: 280, loss: 0.15933075547218323
step: 290, loss: 0.006579921115189791
step: 300, loss: 0.07477002590894699
step: 310, loss: 0.08952976018190384
step: 320, loss: 0.025979988276958466
step: 330, loss: 0.022036032751202583
step: 340, loss: 0.024346599355340004
step: 350, loss: 0.004460613708943129
step: 360, loss: 0.02542649768292904
step: 370, loss: 0.004731003660708666
step: 380, loss: 0.01043845433741808
step: 390, loss: 0.02449987642467022
step: 400, loss: 0.028205275535583496
step: 410, loss: 0.022565225139260292
step: 420, loss: 0.05211041867733002
step: 430, loss: 0.05825033038854599
step: 440, loss: 0.015297717414796352
step: 450, loss: 0.079509437084198
step: 460, loss: 0.012896551750600338
step: 470, loss: 0.12256109714508057
step: 480, loss: 0.11611513048410416
step: 490, loss: 0.02863420732319355
step: 500, loss: 0.04084272310137749
step: 510, loss: 0.04952598363161087
step: 520, loss: 0.037877295166254044
step: 530, loss: 0.1645965278148651
/usr/local/lib/python3.7/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=0.9021636876763877, f1=0.9036827195467423,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01569865643978119
step: 10, loss: 0.02407826855778694
step: 20, loss: 0.0020978860557079315
step: 30, loss: 0.018241021782159805
step: 40, loss: 0.007308726664632559
step: 50, loss: 0.01698964089155197
step: 60, loss: 0.02164490334689617
step: 70, loss: 0.037648364901542664
step: 80, loss: 0.07483972609043121

```

```
step: 90, loss: 0.17527659237384796
step: 100, loss: 0.017547691240906715
step: 110, loss: 0.08341483771800995
step: 120, loss: 0.002551440615206957
step: 130, loss: 0.04945339262485504
step: 140, loss: 0.0205307025462389
step: 150, loss: 0.029245935380458832
step: 160, loss: 0.1780402660369873
step: 170, loss: 0.030457429587841034
step: 180, loss: 0.012798111885786057
step: 190, loss: 0.07360578328371048
step: 200, loss: 0.019467752426862717
step: 210, loss: 0.05747215077280998
step: 220, loss: 0.013759045861661434
step: 230, loss: 0.01996162347495556
step: 240, loss: 0.030906252562999725
step: 250, loss: 0.09365851432085037
step: 260, loss: 0.02675936557352543
step: 270, loss: 0.009040475822985172
step: 280, loss: 0.17107969522476196
step: 290, loss: 0.019335594028234482
step: 300, loss: 0.004291221033781767
step: 310, loss: 0.006876432802528143
step: 320, loss: 0.05005520209670067
step: 330, loss: 0.1308414787054062
step: 340, loss: 0.013417341746389866
step: 350, loss: 0.005071138497442007
step: 360, loss: 0.08563249558210373
step: 370, loss: 0.0031824209727346897
step: 380, loss: 0.005355031695216894
step: 390, loss: 0.0004121899255551398
step: 400, loss: 0.005068385042250156
step: 410, loss: 0.04307830333709717
step: 420, loss: 0.02023025043308735
step: 430, loss: 0.1264687329530716
step: 440, loss: 0.011859320104122162
step: 450, loss: 0.006153311114758253
step: 460, loss: 0.11877915263175964
step: 470, loss: 0.005614117253571749
step: 480, loss: 0.014267893508076668
step: 490, loss: 0.00385110080242157
step: 500, loss: 0.004019572399556637
step: 510, loss: 0.006806913297623396
step: 520, loss: 0.12990786135196686
step: 530, loss: 0.05764206498861313
/usr/local/lib/python3.7/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=0.8993916705662143, f1=0.8992974238875878,  
best_f1=0.9066293183940244  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.011798330582678318  
step: 10, loss: 0.07073735445737839  
step: 20, loss: 0.007519317790865898  
step: 30, loss: 0.0010110235307365656  
step: 40, loss: 0.01348496600985527  
step: 50, loss: 0.006399962119758129  
step: 60, loss: 0.021868186071515083  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
step: 70, loss: 0.0022733728401362896  
step: 80, loss: 0.00034528967808000743  
step: 90, loss: 0.001747853122651577  
step: 100, loss: 0.008312135003507137  
step: 110, loss: 0.0006643941160291433  
step: 120, loss: 0.0030046002939343452  
step: 130, loss: 0.0002484039287082851  
step: 140, loss: 0.05401915684342384  
step: 150, loss: 0.0063459607772529125  
step: 160, loss: 0.006590717006474733  
step: 170, loss: 0.023315558210015297  
step: 180, loss: 0.005407615564763546  
step: 190, loss: 0.002349128946661949  
step: 200, loss: 0.0034223925322294235  
step: 210, loss: 0.03231794014573097  
step: 220, loss: 0.012757787480950356  
step: 230, loss: 0.09759166836738586  
step: 240, loss: 0.0031984946690499783  
step: 250, loss: 0.00709828594699502  
step: 260, loss: 0.03274538740515709  
step: 270, loss: 0.008488517254590988  
step: 280, loss: 0.015876678749918938  
step: 290, loss: 0.11906970292329788  
step: 300, loss: 0.015320667065680027  
step: 310, loss: 0.019250456243753433  
step: 320, loss: 0.04634205996990204  
step: 330, loss: 0.0045089442282915115  
step: 340, loss: 0.005897450726479292  
step: 350, loss: 0.00059078837512061  
step: 360, loss: 0.04916940629482269  
step: 370, loss: 0.04331698268651962  
step: 380, loss: 0.0011502584675326943  
step: 390, loss: 0.007951371371746063  
step: 400, loss: 0.015595278702676296  
step: 410, loss: 0.054428793489933014
```

```
step: 420, loss: 0.003203712170943618
step: 430, loss: 0.005780986975878477
step: 440, loss: 0.15745189785957336
step: 450, loss: 0.08995192497968674
step: 460, loss: 0.0013687562895938754
step: 470, loss: 0.027100345119833946
step: 480, loss: 0.0016496196622028947
step: 490, loss: 0.002561695873737335
step: 500, loss: 0.012499107979238033
step: 510, loss: 0.0161886066198349
step: 520, loss: 0.0450095497071743
step: 530, loss: 0.0016278453404083848
/usr/local/lib/python3.7/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=0.898876404494382, f1=0.896195396899953,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001959907589480281
step: 10, loss: 0.007643344812095165
step: 20, loss: 0.07074648141860962
step: 30, loss: 0.001595405163243413
step: 40, loss: 0.007966279983520508
step: 50, loss: 0.001784533727914095
step: 60, loss: 0.0005145350005477667
step: 70, loss: 0.00038838316686451435
step: 80, loss: 0.0022881021723151207
step: 90, loss: 0.00016264714940916747
step: 100, loss: 0.0011520495172590017
step: 110, loss: 0.0030010573100298643
step: 120, loss: 0.0018916113767772913
step: 130, loss: 0.00023864551621954888
step: 140, loss: 0.0029470480512827635
step: 150, loss: 0.002463166369125247
step: 160, loss: 0.00032915148767642677
step: 170, loss: 0.00032755767460912466
step: 180, loss: 0.0002130036591552198
step: 190, loss: 0.001425933325663209
step: 200, loss: 0.0009207528783008456
step: 210, loss: 0.012919827364385128
step: 220, loss: 0.04706556722521782
step: 230, loss: 0.013461830094456673
step: 240, loss: 0.0007222371059469879
step: 250, loss: 0.004101188387721777
step: 260, loss: 0.001528649590909481
step: 270, loss: 0.03106693923473358
```

```

step: 280, loss: 0.005521080456674099
step: 290, loss: 0.000517316279001534
step: 300, loss: 0.022365592420101166
step: 310, loss: 0.002969134598970413
step: 320, loss: 0.0012278460199013352
step: 330, loss: 0.0010945081012323499
step: 340, loss: 0.047897640615701675
step: 350, loss: 0.026404578238725662
step: 360, loss: 0.003624339820817113
step: 370, loss: 0.0024126458447426558
step: 380, loss: 0.0005728491814807057
step: 390, loss: 0.017456619068980217
step: 400, loss: 0.012547628954052925
step: 410, loss: 0.00431144330650568
step: 420, loss: 0.0025217782240360975
step: 430, loss: 0.000702431658282876
step: 440, loss: 0.00018297068891115487
step: 450, loss: 0.000679230666719377
step: 460, loss: 0.0017552358331158757
step: 470, loss: 0.0033835703507065773
step: 480, loss: 0.0009452891536056995
step: 490, loss: 0.02232344262301922
step: 500, loss: 0.012191607616841793
step: 510, loss: 0.0022906665690243244
step: 520, loss: 0.003912837710231543
step: 530, loss: 0.009601489640772343
/usr/local/lib/python3.7/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=0.8541882109617372, f1=0.8520770010131712,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09145091474056244
step: 10, loss: 0.08475705981254578
step: 20, loss: 0.004835125058889389
step: 30, loss: 0.0020312578417360783
step: 40, loss: 0.05934610962867737
step: 50, loss: 0.06025087088346481
step: 60, loss: 0.0013893176801502705
step: 70, loss: 0.005867209285497665
step: 80, loss: 0.059454295784235
step: 90, loss: 0.05280892923474312
step: 100, loss: 0.01409897394478321
step: 110, loss: 0.0011890458408743143
step: 120, loss: 0.0021469821222126484
step: 130, loss: 0.0042966436594724655

```



```

step: 140, loss: 0.00562291219830513
step: 150, loss: 0.003542455844581127
step: 160, loss: 0.0005789091810584068
step: 170, loss: 0.0010493779554963112
step: 180, loss: 0.0008266806835308671
step: 190, loss: 0.0010997112840414047
step: 200, loss: 0.004117715638130903
step: 210, loss: 0.005196452606469393
step: 220, loss: 0.0020161669235676527
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 230, loss: 0.0032784908544272184
step: 240, loss: 0.024753080680966377
step: 250, loss: 0.0006220964132808149
step: 260, loss: 0.0013264778535813093
step: 270, loss: 0.001580013777129352
step: 280, loss: 0.0014601188013330102
step: 290, loss: 0.00017031475726980716
step: 300, loss: 0.032140281051397324
step: 310, loss: 0.005021471995860338
step: 320, loss: 0.05685483291745186
step: 330, loss: 0.0013547746930271387
step: 340, loss: 0.009721129201352596
step: 350, loss: 0.007458092179149389
step: 360, loss: 0.0020222170278429985
step: 370, loss: 0.00040162098594009876
step: 380, loss: 0.0005313162109814584
step: 390, loss: 0.0003322933043818921
step: 400, loss: 0.0006048395880497992
step: 410, loss: 0.006998786702752113
step: 420, loss: 0.0013360281009227037
step: 430, loss: 0.00010918063344433904
step: 440, loss: 0.00037576662725768983
step: 450, loss: 0.00651740375906229
step: 460, loss: 0.00029964192071929574
step: 470, loss: 0.0010573933832347393
step: 480, loss: 0.23866714537143707
step: 490, loss: 0.006987466476857662
step: 500, loss: 0.00026994399377144873
step: 510, loss: 0.002025479916483164
step: 520, loss: 0.0005013575428165495
step: 530, loss: 0.0016456806333735585
/usr/local/lib/python3.7/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=0.8989186647860837, f1=0.894095595126523,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003175064630340785
step: 10, loss: 0.07656294852495193
step: 20, loss: 0.0018700479995459318
step: 30, loss: 0.0002659682941157371
step: 40, loss: 0.0005497136735357344
step: 50, loss: 0.059704627841711044
step: 60, loss: 0.0012866089818999171
step: 70, loss: 0.0003877556009683758
step: 80, loss: 0.0009031263180077076
step: 90, loss: 0.0006365045555867255
step: 100, loss: 0.004810127429664135
step: 110, loss: 0.01987762749195099
step: 120, loss: 0.004281779285520315
step: 130, loss: 0.007865027524530888
step: 140, loss: 0.003303566947579384
step: 150, loss: 0.07572116702795029
step: 160, loss: 0.15529878437519073
step: 170, loss: 0.0037799112033098936
step: 180, loss: 0.00039712764555588365
step: 190, loss: 0.01301513146609068
step: 200, loss: 0.007266638800501823
step: 210, loss: 0.013688507489860058
step: 220, loss: 0.0033951064106076956
step: 230, loss: 0.0025226627476513386
step: 240, loss: 0.000872583594173193
step: 250, loss: 0.0005851843161508441
step: 260, loss: 9.503518958808854e-05
step: 270, loss: 0.002507178345695138
step: 280, loss: 0.02312338724732399
step: 290, loss: 0.001435974147170782
step: 300, loss: 0.0021455527748912573
step: 310, loss: 0.0070272828452289104
step: 320, loss: 0.009729431010782719
step: 330, loss: 0.0004598561208695173
step: 340, loss: 0.00011634513066383079
step: 350, loss: 0.15739281475543976
step: 360, loss: 0.001043965108692646
step: 370, loss: 0.00019938507466576993
step: 380, loss: 0.001097705913707614
step: 390, loss: 0.05767018720507622
step: 400, loss: 0.08273495733737946
step: 410, loss: 0.0004439139738678932
step: 420, loss: 0.0010127901332452893
step: 430, loss: 0.045647647231817245
step: 440, loss: 0.008268558420240879
step: 450, loss: 9.90795815596357e-05
step: 460, loss: 0.00043708982411772013
```

```

step: 470, loss: 0.0013584191910922527
step: 480, loss: 0.0007485741516575217
step: 490, loss: 0.0530405268073082
step: 500, loss: 0.0005017299554310739
step: 510, loss: 0.014974415302276611
step: 520, loss: 0.0006510508246719837
step: 530, loss: 0.0030609623063355684
/usr/local/lib/python3.7/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=0.9012915129151292, f1=0.9052924791086351,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00501625519245863
step: 10, loss: 0.00036775748594664037
step: 20, loss: 0.039056286215782166
step: 30, loss: 0.00038978736847639084
step: 40, loss: 8.226965292124078e-05
step: 50, loss: 0.004490026738494635
step: 60, loss: 6.833951192675158e-05
step: 70, loss: 0.0029701213352382183
step: 80, loss: 0.028736256062984467
step: 90, loss: 0.001168648130260408
step: 100, loss: 0.0002962926519103348
step: 110, loss: 0.00022495249868370593
step: 120, loss: 0.00023061968386173248
step: 130, loss: 0.0003655939071904868
step: 140, loss: 0.00039549346547573805
step: 150, loss: 0.0038824095390737057
step: 160, loss: 0.0004581470275297761
step: 170, loss: 0.08100370317697525
step: 180, loss: 0.006456985138356686
step: 190, loss: 8.606123446952552e-05
step: 200, loss: 0.0005797009216621518
step: 210, loss: 0.0003431506047490984
step: 220, loss: 0.0019869026727974415
step: 230, loss: 0.00021235542953945696
step: 240, loss: 0.0022145239636301994
step: 250, loss: 7.638331589987502e-05
step: 260, loss: 0.000409450352890417
step: 270, loss: 4.488354898057878e-05
step: 280, loss: 8.870924648363143e-05
step: 290, loss: 0.061984747648239136
step: 300, loss: 0.00012371756020002067
step: 310, loss: 0.007231553550809622
step: 320, loss: 5.434102058643475e-05

```

```

step: 330, loss: 0.025845970958471298
step: 340, loss: 0.004931510426104069
step: 350, loss: 0.00032660941360518336
step: 360, loss: 0.002523049945011735
step: 370, loss: 0.0008061649859882891
step: 380, loss: 0.001140698790550232
step: 390, loss: 8.86790658114478e-05
step: 400, loss: 0.023591380566358566
step: 410, loss: 0.0006939383456483483
step: 420, loss: 0.0003659942594822496
step: 430, loss: 0.0005784990498796105
step: 440, loss: 0.00023689029330853373
step: 450, loss: 9.755208884598687e-05
step: 460, loss: 0.0009722936665639281
step: 470, loss: 3.1578372727381065e-05
step: 480, loss: 6.819577538408339e-05
step: 490, loss: 0.012271403335034847
step: 500, loss: 0.02020960859954357
step: 510, loss: 0.00012393403449095786
step: 520, loss: 0.00024652061983942986
step: 530, loss: 0.0005308522959239781
/usr/local/lib/python3.7/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=0.8915662650602411, f1=0.8802714493456132,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015636080934200436
step: 10, loss: 7.949389691930264e-05
step: 20, loss: 0.0009490328375250101
step: 30, loss: 0.017878249287605286
step: 40, loss: 0.004088717978447676
step: 50, loss: 0.000215407126233913
step: 60, loss: 0.000986688886769116
step: 70, loss: 8.00763227744028e-05
step: 80, loss: 4.712292866315693e-05
step: 90, loss: 0.0016217863885685802
step: 100, loss: 0.0001920322683872655
step: 110, loss: 0.00010787496285047382
step: 120, loss: 0.00010749555804068223
step: 130, loss: 0.00017089811444748193
step: 140, loss: 0.006379060447216034
step: 150, loss: 7.895252201706171e-05
step: 160, loss: 0.00016345951007679105
step: 170, loss: 0.00033791293390095234
step: 180, loss: 0.0001945232506841421

```

```

step: 190, loss: 0.00010388327791588381
step: 200, loss: 0.0003817190299741924
step: 210, loss: 7.514363824157044e-05
step: 220, loss: 0.030531981959939003
step: 230, loss: 0.0005212530377320945
step: 240, loss: 5.4923759307712317e-05
step: 250, loss: 0.00012249792052898556
step: 260, loss: 0.0010863076895475388
step: 270, loss: 6.416703399736434e-05
step: 280, loss: 0.00011160477879457176
step: 290, loss: 0.0018878518603742123
step: 300, loss: 4.669942063628696e-05
step: 310, loss: 0.0015663005178794265
step: 320, loss: 0.006416867021471262
step: 330, loss: 0.00029212544905021787
step: 340, loss: 0.04264157637953758
step: 350, loss: 0.009246254339814186
step: 360, loss: 5.931714258622378e-05
step: 370, loss: 0.0014182627201080322
step: 380, loss: 0.002640998689457774
step: 390, loss: 7.824190106475726e-05
step: 400, loss: 0.00014200800796970725
step: 410, loss: 0.0003399798006284982
step: 420, loss: 0.001327341073192656
step: 430, loss: 5.199392398935743e-05
step: 440, loss: 0.01709640398621559
step: 450, loss: 0.0022299133706837893
step: 460, loss: 8.691888797329739e-05
step: 470, loss: 0.003747442504391074
step: 480, loss: 0.0015775990905240178
step: 490, loss: 0.010049731470644474
step: 500, loss: 0.04678747430443764
step: 510, loss: 4.715202157967724e-05
step: 520, loss: 0.004061440005898476
step: 530, loss: 9.490254160482436e-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=0.9005186232909006, f1=0.8988235294117647,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01327164564281702
step: 10, loss: 0.030335238203406334
step: 20, loss: 0.0007852581329643726
step: 30, loss: 0.0016033223364502192
step: 40, loss: 0.000430736574344337

```

step: 50, loss: 0.00016986274567898363
step: 60, loss: 0.0025930367410182953
step: 70, loss: 0.0015325904823839664
step: 80, loss: 0.00011878336226800457
step: 90, loss: 0.00038309700903482735
step: 100, loss: 0.002825892996042967
step: 110, loss: 0.000887469737790525
step: 120, loss: 0.005827693734318018
step: 130, loss: 4.938096026307903e-05
step: 140, loss: 0.0014163721352815628
step: 150, loss: 3.573416324798018e-05
step: 160, loss: 4.329586954554543e-05
step: 170, loss: 4.523218376561999e-05
step: 180, loss: 5.6970009609358385e-05
step: 190, loss: 0.0019171107560396194
step: 200, loss: 0.006851543672382832
step: 210, loss: 0.0007474057492800057
step: 220, loss: 3.219233258278109e-05
step: 230, loss: 1.6510261048097163e-05
step: 240, loss: 0.0004046458634547889
step: 250, loss: 1.9237102605984546e-05
step: 260, loss: 2.712288187467493e-05
step: 270, loss: 6.470482185250148e-05
step: 280, loss: 0.00024020556884352118
step: 290, loss: 0.0017045480199158192
step: 300, loss: 0.00043674150947481394
step: 310, loss: 0.014366794377565384
step: 320, loss: 4.613984128809534e-05
step: 330, loss: 0.0007627158774994314
step: 340, loss: 0.00024522122112102807
step: 350, loss: 0.0007912563160061836
step: 360, loss: 0.0001851747219916433
step: 370, loss: 8.397374767810106e-05
step: 380, loss: 9.053134999703616e-05
step: 390, loss: 4.76355453429278e-05
step: 400, loss: 2.8948164981557056e-05
step: 410, loss: 8.90292867552489e-05
step: 420, loss: 9.662473166827112e-05
step: 430, loss: 3.338263195473701e-05
step: 440, loss: 4.789338345290162e-05
step: 450, loss: 0.0005343133234418929
step: 460, loss: 0.031209733337163925
step: 470, loss: 0.0026642591692507267
step: 480, loss: 0.00028405693592503667
step: 490, loss: 0.012600107118487358
step: 500, loss: 0.004703808110207319
step: 510, loss: 0.00011982820433331653
step: 520, loss: 3.6715133319376037e-05

```

step: 530, loss: 0.0025844057090580463
/usr/local/lib/python3.7/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=0.8953817153628651, f1=0.890151515151515,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005056214286014438
step: 10, loss: 4.287867704988457e-05
step: 20, loss: 0.00038193902582861483
step: 30, loss: 9.381523705087602e-05
step: 40, loss: 4.236413587932475e-05
step: 50, loss: 0.3565235435962677
step: 60, loss: 0.007412869017571211
step: 70, loss: 0.0012337646912783384
step: 80, loss: 9.463007882004604e-05
step: 90, loss: 0.0012592726852744818
step: 100, loss: 0.002615768928080797
step: 110, loss: 0.0042558168061077595
step: 120, loss: 4.101686499780044e-05
step: 130, loss: 0.00024924284662120044
step: 140, loss: 0.00044293838436715305
step: 150, loss: 0.00024821035913191736
step: 160, loss: 3.822695725830272e-05
step: 170, loss: 5.641002644551918e-05
step: 180, loss: 6.192300497787073e-05
step: 190, loss: 8.888058073353022e-05
step: 200, loss: 0.00021884430316276848
step: 210, loss: 9.066741040442139e-05
step: 220, loss: 0.030441412702202797
step: 230, loss: 0.0010508791310712695
step: 240, loss: 0.0001673347142059356
step: 250, loss: 8.694224379723892e-05
step: 260, loss: 0.00016205149586312473
step: 270, loss: 0.0005325032980181277
step: 280, loss: 0.00012702587991952896
step: 290, loss: 0.0006102841580286622
step: 300, loss: 0.00010946739348582923
step: 310, loss: 0.0010556766064837575
step: 320, loss: 5.34609098394867e-05
step: 330, loss: 0.0001773963449522853
step: 340, loss: 9.471151133766398e-05
step: 350, loss: 0.003797483630478382
step: 360, loss: 0.0030575653072446585
step: 370, loss: 0.012453199364244938
step: 380, loss: 9.818193211685866e-05

```

```

step: 390, loss: 0.00013218987442087382
step: 400, loss: 0.011393913067877293
step: 410, loss: 0.000166424666531384
step: 420, loss: 0.0011122371070086956
step: 430, loss: 0.01801755651831627
step: 440, loss: 0.008350236341357231
step: 450, loss: 5.235791832092218e-05
step: 460, loss: 0.00019813820836134255
step: 470, loss: 6.778456008760259e-05
step: 480, loss: 0.00019752349180635065
step: 490, loss: 0.000591915100812912
step: 500, loss: 4.00980279664509e-05
step: 510, loss: 0.00030786808929406106
step: 520, loss: 3.480043960735202e-05
step: 530, loss: 4.2797426431206986e-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=0.8964539007092198, f1=0.8903846153846154,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002432781271636486
step: 10, loss: 3.447984272497706e-05
step: 20, loss: 0.0002870759926736355
step: 30, loss: 4.1455561586190015e-05
step: 40, loss: 0.00012193768634460866
step: 50, loss: 0.0014636816922575235
step: 60, loss: 0.0002434065390843898
step: 70, loss: 0.00011582157458178699
step: 80, loss: 0.010705910623073578
step: 90, loss: 0.0008913776255212724
step: 100, loss: 0.0007116153719834983
step: 110, loss: 2.0481349565670826e-05
step: 120, loss: 3.6275319871492684e-05
step: 130, loss: 3.253156683058478e-05
step: 140, loss: 7.201225525932387e-05
step: 150, loss: 0.0004496315377764404
step: 160, loss: 0.00018179393373429775
step: 170, loss: 4.1224069718737155e-05
step: 180, loss: 0.00025910887052305043
step: 190, loss: 0.0001338042493443936
step: 200, loss: 0.00023882219102233648
step: 210, loss: 2.4810027753119357e-05
step: 220, loss: 7.270154310390353e-05
step: 230, loss: 3.137303428957239e-05
step: 240, loss: 0.00045767752453684807

```



```

step: 250, loss: 1.7590640709386207e-05
step: 260, loss: 5.6267152103828266e-05
step: 270, loss: 0.00013277748075779527
step: 280, loss: 3.143260255455971e-05
step: 290, loss: 0.0008612123201601207
step: 300, loss: 0.00013304836465977132
step: 310, loss: 2.4325630874955095e-05
step: 320, loss: 0.005312405992299318
step: 330, loss: 3.352021667524241e-05
step: 340, loss: 0.00017954486247617751
step: 350, loss: 0.0007530570146627724
step: 360, loss: 1.7817848856793717e-05
step: 370, loss: 2.904489156208001e-05
step: 380, loss: 2.9417495170491748e-05
step: 390, loss: 0.0004358927544672042
step: 400, loss: 6.358593964250758e-05
step: 410, loss: 9.33703122427687e-05
step: 420, loss: 2.023165870923549e-05
step: 430, loss: 0.001196324359625578
step: 440, loss: 0.00015252636512741446
step: 450, loss: 3.2380899938289076e-05
step: 460, loss: 7.21164105925709e-05
step: 470, loss: 5.290356057230383e-05
step: 480, loss: 0.014952571131289005
step: 490, loss: 3.580906195566058e-05
step: 500, loss: 0.00016491211135871708
step: 510, loss: 0.006014125421643257
step: 520, loss: 0.0025488724932074547
step: 530, loss: 0.00012445647735148668
/usr/local/lib/python3.7/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=0.9052924791086351, f1=0.8992537313432837,
best_f1=0.9066293183940244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003875821130350232
step: 10, loss: 1.8838287360267714e-05
step: 20, loss: 2.0198169295326807e-05
step: 30, loss: 8.832291496219113e-05
step: 40, loss: 0.0007861738558858633
step: 50, loss: 0.0007583448314107955
step: 60, loss: 7.046671089483425e-05
step: 70, loss: 2.4109009245876223e-05
step: 80, loss: 0.000244210590608418
step: 90, loss: 2.7554682674235664e-05
step: 100, loss: 4.408716267789714e-05

```

```
step: 110, loss: 3.048144390049856e-05
step: 120, loss: 3.0284914828371257e-05
step: 130, loss: 0.0006550800753757358
step: 140, loss: 4.152144902036525e-05
step: 150, loss: 1.4327320059237536e-05
step: 160, loss: 1.792212606233079e-05
step: 170, loss: 2.39602813962847e-05
step: 180, loss: 5.9642046835506335e-05
step: 190, loss: 0.00014303527132142335
step: 200, loss: 3.7958780012559146e-05
step: 210, loss: 2.8984628443140537e-05
step: 220, loss: 2.0585253878380172e-05
step: 230, loss: 3.150283373543061e-05
step: 240, loss: 1.5519399312324822e-05
step: 250, loss: 5.618557770503685e-05
step: 260, loss: 2.603129905764945e-05
step: 270, loss: 2.263429996673949e-05
step: 280, loss: 2.010503703786526e-05
step: 290, loss: 0.00472570676356554
step: 300, loss: 5.717786916648038e-05
step: 310, loss: 0.00014723441563546658
step: 320, loss: 3.374897642061114e-05
step: 330, loss: 0.0026414317544549704
step: 340, loss: 2.5729535991558805e-05
step: 350, loss: 0.00012428409536369145
step: 360, loss: 0.015233295038342476
step: 370, loss: 8.200009324355051e-05
step: 380, loss: 2.3696067728451453e-05
step: 390, loss: 0.01980794221162796
step: 400, loss: 2.312234755663667e-05
step: 410, loss: 1.814558527257759e-05
step: 420, loss: 0.000753088213969022
step: 430, loss: 0.00031011548708193004
step: 440, loss: 0.0003940566093660891
step: 450, loss: 5.906872684136033e-05
step: 460, loss: 4.560175511869602e-05
step: 470, loss: 0.00014114684017840773
step: 480, loss: 2.6507465008762665e-05
step: 490, loss: 2.4388447855017148e-05
step: 500, loss: 2.0078994566574693e-05
step: 510, loss: 0.00022869862732477486
step: 520, loss: 1.9676421288750134e-05
step: 530, loss: 3.945023854612373e-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=0.904363974001857, f1=0.8987400839944004,
best_f1=0.9066293183940244
```

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

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

step: 0, loss: 2.177022906835191e-05
step: 10, loss: 2.593050703580957e-05
step: 20, loss: 2.659763595147524e-05
step: 30, loss: 0.0009011070942506194
step: 40, loss: 0.014456008560955524
step: 50, loss: 2.7219904950470664e-05
step: 60, loss: 3.4080410841852427e-05
step: 70, loss: 7.014318543951958e-05
step: 80, loss: 3.5510554880602285e-05
step: 90, loss: 6.584092625416815e-05
step: 100, loss: 6.442212907131761e-05
step: 110, loss: 0.0008479120442643762
step: 120, loss: 0.00033927123877219856
step: 130, loss: 0.024679917842149734
step: 140, loss: 2.682857302716002e-05
step: 150, loss: 3.575652954168618e-05
step: 160, loss: 3.217667472199537e-05
step: 170, loss: 4.902291402686387e-05
step: 180, loss: 1.6592146494076587e-05
step: 190, loss: 0.0005764120724052191
step: 200, loss: 2.634759403008502e-05
step: 210, loss: 0.0005347754340618849
step: 220, loss: 1.8778859157464467e-05
step: 230, loss: 2.4578597731306218e-05
step: 240, loss: 0.002159703290089965
step: 250, loss: 6.8148277932778e-05
step: 260, loss: 1.8406333765597083e-05
step: 270, loss: 2.9726155844400637e-05
step: 280, loss: 4.023977089673281e-05
step: 290, loss: 2.5486209779046476e-05
step: 300, loss: 2.01794937311206e-05
step: 310, loss: 6.372393545461819e-05
step: 320, loss: 0.00035158140235580504
step: 330, loss: 3.788737740251236e-05
step: 340, loss: 1.7478821973782033e-05
step: 350, loss: 2.9471399102476425e-05
step: 360, loss: 0.0007416097796522081
step: 370, loss: 1.4412992641155142e-05
step: 380, loss: 4.364753476693295e-05
step: 390, loss: 0.0016785531770437956
step: 400, loss: 3.6642679333454e-05
step: 410, loss: 2.182223579438869e-05
step: 420, loss: 3.006690167239867e-05
step: 430, loss: 2.5808327336562797e-05
step: 440, loss: 2.3896282073110342e-05

```

step: 450, loss: 1.5757776054670103e-05
step: 460, loss: 1.849207910709083e-05
step: 470, loss: 7.281827129190788e-05
step: 480, loss: 1.5157985217228997e-05
step: 490, loss: 2.6366371457697824e-05
step: 500, loss: 0.0010703125735744834
step: 510, loss: 2.494742329872679e-05
step: 520, loss: 2.4813545678625815e-05
step: 530, loss: 2.035455690929666e-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=0.9020159399906235, f1=0.8962264150943396,
best_f1=0.9066293183940244
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:25, 227.25it/s]
load_f1 = 0.9065550906555091
real_f1 = 0.901109057301294
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 243.61it/s]
```

1.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5516156554222107
step: 10, loss: 0.38477006554603577
step: 20, loss: 0.38197559118270874
step: 30, loss: 0.30619603395462036
step: 40, loss: 0.20554213225841522
step: 50, loss: 0.4644105136394501
step: 60, loss: 0.3182171881198883
step: 70, loss: 0.2343844473361969
step: 80, loss: 0.17086917161941528
step: 90, loss: 0.4773711860179901
step: 100, loss: 0.49917611479759216
step: 110, loss: 0.3038064241409302
step: 120, loss: 0.2667301297187805
step: 130, loss: 0.3832992911338806
step: 140, loss: 0.22242212295532227
step: 150, loss: 0.35010725259780884
step: 160, loss: 0.2899315655231476
step: 170, loss: 0.32188138365745544
step: 180, loss: 0.15375567972660065
step: 190, loss: 0.22756078839302063
step: 200, loss: 0.34526899456977844

```

```

step: 210, loss: 0.1913747936487198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3719745222929936, f1=0.39800995024875624,
best_f1=0.39800995024875624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1533277928829193
step: 10, loss: 0.3083803057670593
step: 20, loss: 0.23382870852947235
step: 30, loss: 0.28334102034568787
step: 40, loss: 0.18506555259227753
step: 50, loss: 0.11894641071557999
step: 60, loss: 0.4237271845340729
step: 70, loss: 0.1399688720703125
step: 80, loss: 0.3485444486141205
step: 90, loss: 0.10670174658298492
step: 100, loss: 0.03337734937667847
step: 110, loss: 0.1745401918888092
step: 120, loss: 0.12911127507686615
step: 130, loss: 0.05666111037135124
step: 140, loss: 0.21454600989818573
step: 150, loss: 0.27788597345352173
step: 160, loss: 0.24057255685329437
step: 170, loss: 0.18898767232894897
step: 180, loss: 0.2926322817802429
step: 190, loss: 0.2771397829055786
step: 200, loss: 0.08542478084564209
step: 210, loss: 0.2429238259792328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.48394004282655245, f1=0.4948024948024948,
best_f1=0.4948024948024948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11109201610088348

```

```

step: 10, loss: 0.2594306170940399
step: 20, loss: 0.19941440224647522
step: 30, loss: 0.22817020118236542
step: 40, loss: 0.10872134566307068
step: 50, loss: 0.12735715508460999
step: 60, loss: 0.18853704631328583
step: 70, loss: 0.3448071777820587
step: 80, loss: 0.09647605568170547
step: 90, loss: 0.13641822338104248
step: 100, loss: 0.19696944952011108
step: 110, loss: 0.22750389575958252
step: 120, loss: 0.17160575091838837
step: 130, loss: 0.17155785858631134
step: 140, loss: 0.10488978028297424
step: 150, loss: 0.27466318011283875
step: 160, loss: 0.05423179641366005
step: 170, loss: 0.19240456819534302
step: 180, loss: 0.17687398195266724
step: 190, loss: 0.22847431898117065
step: 200, loss: 0.13052307069301605
step: 210, loss: 0.07549970597028732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.499089253187614, f1=0.5261194029850746,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2525823414325714
step: 10, loss: 0.24870768189430237
step: 20, loss: 0.12024005502462387
step: 30, loss: 0.12344493716955185
step: 40, loss: 0.061270032078027725
step: 50, loss: 0.15474426746368408
step: 60, loss: 0.3138546049594879
step: 70, loss: 0.23443648219108582
step: 80, loss: 0.11089975386857986
step: 90, loss: 0.07817861437797546
step: 100, loss: 0.32061752676963806
step: 110, loss: 0.33689695596694946
step: 120, loss: 0.17507876455783844
step: 130, loss: 0.0988142117857933
step: 140, loss: 0.2725219428539276

```



```

step: 150, loss: 0.17009393870830536
step: 160, loss: 0.06594906002283096
step: 170, loss: 0.3469376862049103
step: 180, loss: 0.4904560446739197
step: 190, loss: 0.19473664462566376
step: 200, loss: 0.22540441155433655
step: 210, loss: 0.2776365280151367
/usr/local/lib/python3.7/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=0.4827586206896552, f1=0.4910394265232975,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2569030523300171
step: 10, loss: 0.18254971504211426
step: 20, loss: 0.24017348885536194
step: 30, loss: 0.134999617934227
step: 40, loss: 0.0593431331217289
step: 50, loss: 0.3756142854690552
step: 60, loss: 0.0885249450802803
step: 70, loss: 0.08782108128070831
step: 80, loss: 0.0756634920835495
step: 90, loss: 0.13494233787059784
step: 100, loss: 0.007733840495347977
step: 110, loss: 0.2757951021194458
step: 120, loss: 0.19189248979091644
step: 130, loss: 0.15138082206249237
step: 140, loss: 0.22149546444416046
step: 150, loss: 0.10888627171516418
step: 160, loss: 0.09647073596715927
step: 170, loss: 0.14804446697235107
step: 180, loss: 0.04784199595451355
step: 190, loss: 0.04498615860939026
step: 200, loss: 0.23579128086566925
step: 210, loss: 0.025125056505203247
/usr/local/lib/python3.7/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=0.4868421052631579, f1=0.47511312217194573,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06880190223455429
step: 10, loss: 0.12586544454097748
step: 20, loss: 0.05925082415342331

```

```

step: 30, loss: 0.006538929417729378
step: 40, loss: 0.06669461727142334
step: 50, loss: 0.04280814155936241
step: 60, loss: 0.14741846919059753
step: 70, loss: 0.009596450254321098
step: 80, loss: 0.06302373856306076
step: 90, loss: 0.4166621267795563
step: 100, loss: 0.026417803019285202
step: 110, loss: 0.10127130150794983
step: 120, loss: 0.024881117045879364
step: 130, loss: 0.19177110493183136
step: 140, loss: 0.08708412945270538
step: 150, loss: 0.056750085204839706
step: 160, loss: 0.024912571534514427
step: 170, loss: 0.1674778312444687
step: 180, loss: 0.047472018748521805
step: 190, loss: 0.07525449991226196
step: 200, loss: 0.05617377907037735
step: 210, loss: 0.05711009353399277
/usr/local/lib/python3.7/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=0.46393762183235865, f1=0.4675834970530452,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007847190834581852
step: 10, loss: 0.08267977833747864
step: 20, loss: 0.02713206224143505
step: 30, loss: 0.02019021101295948
step: 40, loss: 0.07252655178308487
step: 50, loss: 0.15607735514640808
step: 60, loss: 0.06168413907289505
step: 70, loss: 0.027273310348391533
step: 80, loss: 0.08853266388177872
step: 90, loss: 0.05655086040496826
step: 100, loss: 0.029626768082380295
step: 110, loss: 0.16982360184192657
step: 120, loss: 0.03361399844288826
step: 130, loss: 0.1209927499294281
step: 140, loss: 0.04044698178768158
step: 150, loss: 0.027437124401330948
step: 160, loss: 0.042170606553554535
step: 170, loss: 0.04317723214626312
step: 180, loss: 0.02494535781443119
step: 190, loss: 0.09734640270471573
step: 200, loss: 0.052050478756427765

```

```

step: 210, loss: 0.11092385649681091
/usr/local/lib/python3.7/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=0.4496124031007752, f1=0.4591439688715953,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003918835427612066
step: 10, loss: 0.1395975798368454
step: 20, loss: 0.017973007634282112
step: 30, loss: 0.029602685943245888
step: 40, loss: 0.05412667989730835
step: 50, loss: 0.10044778138399124
step: 60, loss: 0.14893142879009247
step: 70, loss: 0.018060797825455666
step: 80, loss: 0.10123546421527863
step: 90, loss: 0.008479571901261806
step: 100, loss: 0.04757295548915863
step: 110, loss: 0.14958034455776215
step: 120, loss: 0.015084651298820972
step: 130, loss: 0.004053195007145405
step: 140, loss: 0.010944307781755924
step: 150, loss: 0.11694692820310593
step: 160, loss: 0.016283133998513222
step: 170, loss: 0.015041464939713478
step: 180, loss: 0.040019307285547256
step: 190, loss: 0.00621245289221406
step: 200, loss: 0.0023442988749593496
step: 210, loss: 0.11992128193378448
/usr/local/lib/python3.7/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=0.47418738049713194, f1=0.46062992125984253,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04081298038363457
step: 10, loss: 0.04796475172042847
step: 20, loss: 0.010475976392626762
step: 30, loss: 0.01630445383489132
step: 40, loss: 0.2170509546995163
step: 50, loss: 0.17539246380329132
step: 60, loss: 0.13735075294971466
step: 70, loss: 0.04289231821894646
step: 80, loss: 0.017688343301415443

```

```

step: 90, loss: 0.042578890919685364
step: 100, loss: 0.02252950333058834
step: 110, loss: 0.10936256498098373
step: 120, loss: 0.00451258011162281
step: 130, loss: 0.059755075722932816
step: 140, loss: 0.02665117383003235
step: 150, loss: 0.26198381185531616
step: 160, loss: 0.01186465099453926
step: 170, loss: 0.04049847647547722
step: 180, loss: 0.03276890143752098
step: 190, loss: 0.044459857046604156
step: 200, loss: 0.038454800844192505
step: 210, loss: 0.01773499697446823
/usr/local/lib/python3.7/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=0.44181459566074954, f1=0.4572564612326043,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13565416634082794
step: 10, loss: 0.0369541272521019
step: 20, loss: 0.001898337621241808
step: 30, loss: 0.1308075487613678
step: 40, loss: 0.006459391675889492
step: 50, loss: 0.07951418310403824
step: 60, loss: 0.009913794696331024
step: 70, loss: 0.03296234831213951
step: 80, loss: 0.02335839346051216
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 90, loss: 0.07653330266475677
step: 100, loss: 0.0036351759918034077
step: 110, loss: 0.06344561278820038
step: 120, loss: 0.012345103546977043
step: 130, loss: 0.0573861338198185
step: 140, loss: 0.0023903357796370983
step: 150, loss: 0.02234419248998165
step: 160, loss: 0.11647326499223709
step: 170, loss: 0.01529106218367815
step: 180, loss: 0.09432677924633026
step: 190, loss: 0.2618560492992401
step: 200, loss: 0.20962022244930267
step: 210, loss: 0.09123227000236511
/usr/local/lib/python3.7/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=0.444043321299639, f1=0.45472061657032753,

```

```

best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02646823786199093
step: 10, loss: 0.005863533820956945
step: 20, loss: 0.01855272240936756
step: 30, loss: 0.013707274571061134
step: 40, loss: 0.025727737694978714
step: 50, loss: 0.004527258221060038
step: 60, loss: 0.18236221373081207
step: 70, loss: 0.007959089241921902
step: 80, loss: 0.2354428768157959
step: 90, loss: 0.045560263097286224
step: 100, loss: 0.02423195354640484
step: 110, loss: 0.13451717793941498
step: 120, loss: 0.00803031399846077
step: 130, loss: 0.012447762303054333
step: 140, loss: 0.015983447432518005
step: 150, loss: 0.0039151571691036224
step: 160, loss: 0.004867636598646641
step: 170, loss: 0.006516020279377699
step: 180, loss: 0.003499542595818639
step: 190, loss: 0.02299739606678486
step: 200, loss: 0.014023983851075172
step: 210, loss: 0.004483662545681
/usr/local/lib/python3.7/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=0.4618181818181818, f1=0.4635514018691589,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008174590766429901
step: 10, loss: 0.0017432703170925379
step: 20, loss: 0.03356987237930298
step: 30, loss: 0.07889938354492188
step: 40, loss: 0.10811200737953186
step: 50, loss: 0.0015757700894027948
step: 60, loss: 0.10353279858827591
step: 70, loss: 0.00498080113902688
step: 80, loss: 0.0043104724027216434
step: 90, loss: 0.009373034350574017
step: 100, loss: 0.022868363186717033
step: 110, loss: 0.009655038826167583
step: 120, loss: 0.027858102694153786
step: 130, loss: 0.004167599603533745

```

```

step: 140, loss: 0.003110239515081048
step: 150, loss: 0.0438871793448925
step: 160, loss: 0.004876322578638792
step: 170, loss: 0.0039971135556697845
step: 180, loss: 0.033792782574892044
step: 190, loss: 0.024890802800655365
step: 200, loss: 0.012060887180268764
step: 210, loss: 0.0051330639980733395
/usr/local/lib/python3.7/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=0.4410646387832699, f1=0.45999999999999996,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022479195147752762
step: 10, loss: 0.004918100778013468
step: 20, loss: 0.03411475196480751
step: 30, loss: 0.05569113790988922
step: 40, loss: 0.0038465680554509163
step: 50, loss: 0.025729266926646233
step: 60, loss: 0.004916914738714695
step: 70, loss: 0.06199216470122337
step: 80, loss: 0.006312616169452667
step: 90, loss: 0.0005969438934698701
step: 100, loss: 0.0027393994387239218
step: 110, loss: 0.0024386662989854813
step: 120, loss: 0.17738182842731476
step: 130, loss: 0.00345748127438128
step: 140, loss: 0.001231921254657209
step: 150, loss: 0.010775333270430565
step: 160, loss: 0.002273445250466466
step: 170, loss: 0.012862518429756165
step: 180, loss: 0.22868947684764862
step: 190, loss: 0.0006387735484167933
step: 200, loss: 0.019944507628679276
step: 210, loss: 0.016583522781729698
/usr/local/lib/python3.7/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=0.4555984555984556, f1=0.4476386036960986,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016045838128775358
step: 10, loss: 0.037996575236320496

```

```

step: 20, loss: 0.007284677121788263
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.011354937218129635
step: 40, loss: 0.002710650209337473
step: 50, loss: 0.004760422743856907
step: 60, loss: 0.025784680619835854
step: 70, loss: 0.0064803543500602245
step: 80, loss: 0.003685828996822238
step: 90, loss: 0.10225604474544525
step: 100, loss: 0.13768118619918823
step: 110, loss: 0.004294858779758215
step: 120, loss: 0.0029857198242098093
step: 130, loss: 0.013033991679549217
step: 140, loss: 0.07320744544267654
step: 150, loss: 0.014955899678170681
step: 160, loss: 0.004324646200984716
step: 170, loss: 0.13887271285057068
step: 180, loss: 0.00494109932333231
step: 190, loss: 0.02359004318714142
step: 200, loss: 0.007121098227798939
step: 210, loss: 0.004970721900463104
/usr/local/lib/python3.7/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=0.44799999999999995, f1=0.45473684210526316,
best_f1=0.5261194029850746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024806340225040913
step: 10, loss: 0.005002561490982771
step: 20, loss: 0.06693369895219803
step: 30, loss: 0.014626901596784592
step: 40, loss: 0.0014346996322274208
step: 50, loss: 0.012044484727084637
step: 60, loss: 0.08696925640106201
step: 70, loss: 0.0020415170583873987
step: 80, loss: 0.0015838347608223557
step: 90, loss: 0.002966947853565216
step: 100, loss: 0.000835337326861918
step: 110, loss: 0.03868332505226135
step: 120, loss: 0.0031581635121256113
step: 130, loss: 0.0014434022596105933
step: 140, loss: 0.0004986025742255151
step: 150, loss: 0.0009603235521353781
step: 160, loss: 0.015811007469892502
step: 170, loss: 0.0006271664751693606
step: 180, loss: 0.0017444450641050935

```

```

step: 190, loss: 0.006951163988560438
step: 200, loss: 0.0022421462927013636
step: 210, loss: 0.1635756641626358
/usr/local/lib/python3.7/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=0.4488977955911824, f1=0.45702306079664573,
best_f1=0.5261194029850746
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 463.20it/s]
load_f1 = 0.5062388591800357
real_f1 = 0.5
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 269.82it/s]

```


1.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5669991374015808
step: 10, loss: 0.3563770055770874
step: 20, loss: 0.28573769330978394
step: 30, loss: 0.4394654333591461
step: 40, loss: 0.41891783475875854
step: 50, loss: 0.3046737015247345
step: 60, loss: 0.2950672507286072
step: 70, loss: 0.2921471893787384
step: 80, loss: 0.23104652762413025
step: 90, loss: 0.31017622351646423
step: 100, loss: 0.31596043705940247
step: 110, loss: 0.35721296072006226
step: 120, loss: 0.11731131374835968
step: 130, loss: 0.14145933091640472
step: 140, loss: 0.08052177727222443
step: 150, loss: 0.12646172940731049
step: 160, loss: 0.10208643972873688
step: 170, loss: 0.2104695439338684
step: 180, loss: 0.03449787572026253
step: 190, loss: 0.2655513286590576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6849315068493149, f1=0.6770833333333335,
best_f1=0.6770833333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18866702914237976
step: 10, loss: 0.0690009593963623
step: 20, loss: 0.19513562321662903
step: 30, loss: 0.1526561826467514
step: 40, loss: 0.15869739651679993
step: 50, loss: 0.07093410938978195
step: 60, loss: 0.32272598147392273
step: 70, loss: 0.1362772285938263
step: 80, loss: 0.11236918717622757
step: 90, loss: 0.16504065692424774
step: 100, loss: 0.03168398141860962
step: 110, loss: 0.15045712888240814
step: 120, loss: 0.3363989293575287
step: 130, loss: 0.08532950282096863
step: 140, loss: 0.05877063795924187
step: 150, loss: 0.19898805022239685
step: 160, loss: 0.10544648766517639
step: 170, loss: 0.18033096194267273
step: 180, loss: 0.11083055287599564
step: 190, loss: 0.026221629232168198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7324675324675326, f1=0.7593052109181142,
best_f1=0.7593052109181142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.049716413021087646
step: 10, loss: 0.2050304412841797
step: 20, loss: 0.15851512551307678
step: 30, loss: 0.2790665626525879
step: 40, loss: 0.20573388040065765
step: 50, loss: 0.340950608253479
step: 60, loss: 0.08331672102212906
step: 70, loss: 0.09520358592271805
step: 80, loss: 0.11196332424879074
step: 90, loss: 0.05900635942816734

```

```

step: 100, loss: 0.045510560274124146
step: 110, loss: 0.046051446348428726
step: 120, loss: 0.02999313361942768
step: 130, loss: 0.04230271652340889
step: 140, loss: 0.07068148255348206
step: 150, loss: 0.15951426327228546
step: 160, loss: 0.04452424496412277
step: 170, loss: 0.08258774876594543
step: 180, loss: 0.08265149593353271
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.1311793476343155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7712082262210797, f1=0.7719298245614035,
best_f1=0.7719298245614035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02290402166545391
step: 10, loss: 0.22268663346767426
step: 20, loss: 0.01593191549181938
step: 30, loss: 0.05950024351477623
step: 40, loss: 0.15862715244293213
step: 50, loss: 0.09621763229370117
step: 60, loss: 0.11116162687540054
step: 70, loss: 0.13005569577217102
step: 80, loss: 0.17257997393608093
step: 90, loss: 0.026484280824661255
step: 100, loss: 0.10829020291566849
step: 110, loss: 0.006486358121037483
step: 120, loss: 0.10996086150407791
step: 130, loss: 0.15278463065624237
step: 140, loss: 0.11043108999729156
step: 150, loss: 0.08920034766197205
step: 160, loss: 0.05569033324718475
step: 170, loss: 0.07047149538993835
step: 180, loss: 0.03861548751592636
step: 190, loss: 0.13160118460655212
/usr/local/lib/python3.7/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=0.76, f1=0.7648725212464589, best_f1=0.7719298245614035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.05891585350036621
step: 10, loss: 0.005465284921228886
step: 20, loss: 0.05822551250457764
step: 30, loss: 0.011787984520196915
step: 40, loss: 0.05598985403776169
step: 50, loss: 0.17650356888771057
step: 60, loss: 0.033980827778577805
step: 70, loss: 0.0244512390345335
step: 80, loss: 0.0930861085653305
step: 90, loss: 0.01298443228006363
step: 100, loss: 0.002018332714214921
step: 110, loss: 0.01387566328048706
step: 120, loss: 0.09833719581365585
step: 130, loss: 0.012527018785476685
step: 140, loss: 0.16744209825992584
step: 150, loss: 0.0675189271569252
step: 160, loss: 0.024159317836165428
step: 170, loss: 0.009513855911791325
step: 180, loss: 0.13006770610809326
step: 190, loss: 0.10511761903762817
```

```
/usr/local/lib/python3.7/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=0.7653061224489796, f1=0.7193877551020408,
best_f1=0.7719298245614035
```

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

An input tensor was not cuda.

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

```
step: 0, loss: 0.011553010903298855
step: 10, loss: 0.09094428271055222
step: 20, loss: 0.0046874480322003365
step: 30, loss: 0.03724236786365509
step: 40, loss: 0.14281520247459412
step: 50, loss: 0.00857336912304163
step: 60, loss: 0.0712134838104248
step: 70, loss: 0.0340469554066658
step: 80, loss: 0.0516098290681839
step: 90, loss: 0.03014385513961315
step: 100, loss: 0.0016339208232238889
step: 110, loss: 0.019108258187770844
step: 120, loss: 0.2703437805175781
step: 130, loss: 0.09979597479104996
step: 140, loss: 0.02070734277367592
step: 150, loss: 0.011403489857912064
step: 160, loss: 0.05194355547428131
step: 170, loss: 0.037370339035987854
```

```

step: 180, loss: 0.04656781628727913
step: 190, loss: 0.06716009229421616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7734806629834254, f1=0.7826086956521738,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013160320930182934
step: 10, loss: 0.005679173860698938
step: 20, loss: 0.07729094475507736
step: 30, loss: 0.02996532991528511
step: 40, loss: 0.004561900161206722
step: 50, loss: 0.18137557804584503
step: 60, loss: 0.08934424072504044
step: 70, loss: 0.0022073527798056602
step: 80, loss: 0.014499600976705551
step: 90, loss: 0.00247846357524395
step: 100, loss: 0.0028896601870656013
step: 110, loss: 0.012071935459971428
step: 120, loss: 0.0011146850883960724
step: 130, loss: 0.04472050443291664
step: 140, loss: 0.028603294864296913
step: 150, loss: 0.22888098657131195
step: 160, loss: 0.13699595630168915
step: 170, loss: 0.021238449960947037
step: 180, loss: 0.004311729222536087
step: 190, loss: 0.004826403688639402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7811634349030471, f1=0.7845303867403315,
best_f1=0.7845303867403315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004266939125955105
step: 10, loss: 0.006062089465558529

```

```

step: 20, loss: 0.2152368277311325
step: 30, loss: 0.0035333908163011074
step: 40, loss: 0.021324975416064262
step: 50, loss: 0.0011088327737525105
step: 60, loss: 0.003471772186458111
step: 70, loss: 0.02043156325817108
step: 80, loss: 0.0014933552592992783
step: 90, loss: 0.008824458345770836
step: 100, loss: 0.01855318807065487
step: 110, loss: 0.14381645619869232
step: 120, loss: 0.02849585935473442
step: 130, loss: 0.0023078974336385727
step: 140, loss: 0.0972885936498642
step: 150, loss: 0.08836150914430618
step: 160, loss: 0.06466471403837204
step: 170, loss: 0.004194925539195538
step: 180, loss: 0.05116704851388931
step: 190, loss: 0.02970467321574688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.79155672823219, f1=0.772845953002611,
best_f1=0.772845953002611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01914818212389946
step: 10, loss: 0.002476002089679241
step: 20, loss: 0.0023341828491538763
step: 30, loss: 0.026942085474729538
step: 40, loss: 0.0026212024968117476
step: 50, loss: 0.009508755058050156
step: 60, loss: 0.0010891694109886885
step: 70, loss: 0.003264941740781069
step: 80, loss: 0.0012104682391509414
step: 90, loss: 0.004059562459588051
step: 100, loss: 0.0019012034172192216
step: 110, loss: 0.004923532251268625
step: 120, loss: 0.0023412536829710007
step: 130, loss: 0.004196696914732456
step: 140, loss: 0.031893059611320496
step: 150, loss: 0.0015740209491923451
step: 160, loss: 0.0005385020049288869
step: 170, loss: 0.0026676184497773647

```

```

step: 180, loss: 0.001990197692066431
step: 190, loss: 0.0008229318773373961
/usr/local/lib/python3.7/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=0.7838616714697405, f1=0.7392550143266476,
best_f1=0.772845953002611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004064297303557396
step: 10, loss: 0.0007253009243868291
step: 20, loss: 0.010397692210972309
step: 30, loss: 0.00033048729528672993
step: 40, loss: 0.006299017462879419
step: 50, loss: 0.016630878672003746
step: 60, loss: 0.0020211581140756607
step: 70, loss: 0.07199663668870926
step: 80, loss: 0.0009346367442049086
step: 90, loss: 0.003279935335740447
step: 100, loss: 0.008070547133684158
step: 110, loss: 0.05352494493126869
step: 120, loss: 0.08130523562431335
step: 130, loss: 0.0029940688982605934
step: 140, loss: 0.006980027537792921
step: 150, loss: 0.017591720446944237
step: 160, loss: 0.020572267472743988
step: 170, loss: 0.005191023461520672
step: 180, loss: 0.15574753284454346
step: 190, loss: 0.002033549128100276
/usr/local/lib/python3.7/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=0.7880434782608696, f1=0.7870619946091645,
best_f1=0.772845953002611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006669610156677663
step: 10, loss: 0.007736229337751865
step: 20, loss: 0.004706737585365772
step: 30, loss: 0.0038105370476841927
step: 40, loss: 0.003517244243994355
step: 50, loss: 0.005147071089595556
step: 60, loss: 0.026927510276436806
step: 70, loss: 0.0012742863036692142
step: 80, loss: 0.03159359470009804
step: 90, loss: 0.0015011432114988565

```



```

step: 100, loss: 0.00052467524074018
step: 110, loss: 0.0014479487435892224
step: 120, loss: 0.0009463653550483286
step: 130, loss: 0.0027082376182079315
step: 140, loss: 0.006891028489917517
step: 150, loss: 0.0011933700880035758
step: 160, loss: 0.0041369847021996975
step: 170, loss: 0.0006030541262589395
step: 180, loss: 0.013475948013365269
step: 190, loss: 0.001628492260351777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.793201133144476, f1=0.7422096317280454,
best_f1=0.7422096317280454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008152599330060184
step: 10, loss: 0.000783736992161721
step: 20, loss: 0.11330869048833847
step: 30, loss: 0.0018019923008978367
step: 40, loss: 0.0030975118279457092
step: 50, loss: 0.025294341146945953
step: 60, loss: 0.000797385408077389
step: 70, loss: 0.000722641940228641
step: 80, loss: 0.0005850344314239919
step: 90, loss: 0.01746349222958088
step: 100, loss: 0.00044816863373853266
step: 110, loss: 0.0003317016235087067
step: 120, loss: 0.02899681217968464
step: 130, loss: 0.0016329885693266988
step: 140, loss: 0.0019457279704511166
step: 150, loss: 0.00069981295382604
step: 160, loss: 0.0002679827739484608
step: 170, loss: 0.0015019080601632595
step: 180, loss: 0.0005528354668058455
step: 190, loss: 0.0018998049199581146
/usr/local/lib/python3.7/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=0.7786666666666667, f1=0.765625, best_f1=0.7422096317280454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21870063245296478
step: 10, loss: 0.051601167768239975
step: 20, loss: 0.006463070400059223
step: 30, loss: 0.0019322825828567147
step: 40, loss: 0.0005759296473115683
step: 50, loss: 0.005157663021236658
step: 60, loss: 0.0008377806516364217
step: 70, loss: 0.0012029180070385337
step: 80, loss: 0.0010068066185340285
step: 90, loss: 0.0011775654274970293
step: 100, loss: 0.0003628105332609266
step: 110, loss: 0.001372704515233636
step: 120, loss: 0.0008184186881408095
step: 130, loss: 0.0021516787819564342
step: 140, loss: 0.0003866353363264352
step: 150, loss: 0.001415325328707695
step: 160, loss: 0.004119126591831446
step: 170, loss: 0.0004073833115398884
step: 180, loss: 0.0009091104147955775
step: 190, loss: 0.008930196985602379
/usr/local/lib/python3.7/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=0.7735368956743003, f1=0.7551020408163265,
best_f1=0.7422096317280454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022467102389782667
step: 10, loss: 0.0008750377455726266
step: 20, loss: 0.0007370767998509109
step: 30, loss: 0.0002827641146723181
step: 40, loss: 0.004076520446687937
step: 50, loss: 0.0007670930353924632
step: 60, loss: 0.0719413161277771
step: 70, loss: 0.0019373701652511954
step: 80, loss: 0.0009199320338666439
step: 90, loss: 0.0036152570974081755
step: 100, loss: 0.12206050753593445
step: 110, loss: 0.005820742342621088
step: 120, loss: 0.18409976363182068
step: 130, loss: 0.0003632090229075402
step: 140, loss: 0.0026273152325302362
step: 150, loss: 0.0004595224454533309
step: 160, loss: 0.0005326498067006469
step: 170, loss: 0.0013971652369946241
step: 180, loss: 0.027370641008019447

```

```

step: 190, loss: 0.004341017920523882
/usr/local/lib/python3.7/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=0.7777777777777779, f1=0.7506702412868632,
best_f1=0.7422096317280454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033171274699270725
step: 10, loss: 0.0010006467346102
step: 20, loss: 0.0007933027809485793
step: 30, loss: 0.0008430028101429343
step: 40, loss: 0.0005395353073254228
step: 50, loss: 0.01456774678081274
step: 60, loss: 0.0009720121161080897
step: 70, loss: 0.026932979002594948
step: 80, loss: 0.00047225228627212346
step: 90, loss: 0.00037791821523569524
step: 100, loss: 0.0005337080219760537
step: 110, loss: 0.0009265955886803567
step: 120, loss: 0.00040430252556689084
step: 130, loss: 0.0005447946605272591
step: 140, loss: 0.0016626815777271986
step: 150, loss: 0.0009456851985305548
step: 160, loss: 0.0017015996854752302
step: 170, loss: 0.0005457581137306988
step: 180, loss: 0.014649055898189545
step: 190, loss: 0.004592378623783588
/usr/local/lib/python3.7/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=0.7835616438356164, f1=0.7452054794520547,
best_f1=0.7422096317280454
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:08, 239.62it/s]
load_f1 = 0.7365728900255755
real_f1 = 0.7104622871046229
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 271.27it/s]

```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6306028366088867
step: 10, loss: 0.375547856092453
step: 20, loss: 0.2945812940597534
step: 30, loss: 0.3802810609340668
step: 40, loss: 0.2842734158039093
step: 50, loss: 0.2595020830631256
step: 60, loss: 0.24579459428787231
step: 70, loss: 0.3723761737346649
step: 80, loss: 0.3605038523674011
step: 90, loss: 0.2873445451259613
step: 100, loss: 0.17969846725463867
step: 110, loss: 0.2084912806749344
step: 120, loss: 0.2148619294166565
step: 130, loss: 0.06457886099815369
step: 140, loss: 0.3348064124584198
step: 150, loss: 0.39388591051101685
step: 160, loss: 0.1294320821762085
step: 170, loss: 0.20263883471488953
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.684863523573201, f1=0.6634382566585957,
best_f1=0.6634382566585957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.079014852643013
step: 10, loss: 0.19095943868160248
step: 20, loss: 0.04997483268380165
step: 30, loss: 0.17313706874847412
step: 40, loss: 0.0754399448633194
step: 50, loss: 0.18916182219982147
step: 60, loss: 0.3007543683052063
step: 70, loss: 0.08682786673307419
step: 80, loss: 0.08088172227144241
step: 90, loss: 0.1440044343471527
step: 100, loss: 0.14541016519069672
step: 110, loss: 0.07066642493009567
step: 120, loss: 0.20166926085948944
step: 130, loss: 0.02898235246539116
step: 140, loss: 0.24286778271198273
step: 150, loss: 0.1733795702457428

```

```

step: 160, loss: 0.19295963644981384
step: 170, loss: 0.09779797494411469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7427055702917773, f1=0.6961038961038961,
best_f1=0.6961038961038961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10328592360019684
step: 10, loss: 0.06677012890577316
step: 20, loss: 0.0222457367926836
step: 30, loss: 0.07060609757900238
step: 40, loss: 0.07010503113269806
step: 50, loss: 0.17225901782512665
step: 60, loss: 0.27802643179893494
step: 70, loss: 0.16293105483055115
step: 80, loss: 0.0669272318482399
step: 90, loss: 0.07312189042568207
step: 100, loss: 0.02546011656522751
step: 110, loss: 0.14184501767158508
step: 120, loss: 0.07851701229810715
step: 130, loss: 0.0701800137758255
step: 140, loss: 0.07976000756025314
step: 150, loss: 0.02307366020977497
step: 160, loss: 0.08416341245174408
step: 170, loss: 0.0746067464351654
/usr/local/lib/python3.7/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=0.7303370786516854, f1=0.6933333333333334,
best_f1=0.6961038961038961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026310259476304054
step: 10, loss: 0.04835205897688866
step: 20, loss: 0.014103780500590801
step: 30, loss: 0.12760713696479797
step: 40, loss: 0.002269486663863063
step: 50, loss: 0.16953463852405548
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.12996380031108856

```

```

step: 70, loss: 0.034537266939878464
step: 80, loss: 0.08200138807296753
step: 90, loss: 0.044092100113630295
step: 100, loss: 0.13309752941131592
step: 110, loss: 0.07564502954483032
step: 120, loss: 0.031142832711338997
step: 130, loss: 0.19486136734485626
step: 140, loss: 0.041177138686180115
step: 150, loss: 0.04464639350771904
step: 160, loss: 0.03293468803167343
step: 170, loss: 0.05767257511615753
/usr/local/lib/python3.7/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=0.7287234042553191, f1=0.6788511749347259,
best_f1=0.6961038961038961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07500722259283066
step: 10, loss: 0.12597286701202393
step: 20, loss: 0.011574970558285713
step: 30, loss: 0.03641238063573837
step: 40, loss: 0.004987575579434633
step: 50, loss: 0.11135996133089066
step: 60, loss: 0.02249016985297203
step: 70, loss: 0.21048504114151
step: 80, loss: 0.10747247189283371
step: 90, loss: 0.08123087137937546
step: 100, loss: 0.12578809261322021
step: 110, loss: 0.18662677705287933
step: 120, loss: 0.021726766601204872
step: 130, loss: 0.0921863317489624
step: 140, loss: 0.10668592900037766
step: 150, loss: 0.1381109058856964
step: 160, loss: 0.207439124584198
step: 170, loss: 0.21252337098121643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7487179487179487, f1=0.6891566265060242,
best_f1=0.6891566265060242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007019991520792246
step: 10, loss: 0.09705978631973267
step: 20, loss: 0.00596743356436491
step: 30, loss: 0.04155189171433449
step: 40, loss: 0.014034731313586235
step: 50, loss: 0.11586020886898041
step: 60, loss: 0.2002231627702713
step: 70, loss: 0.06904945522546768
step: 80, loss: 0.029152492061257362
step: 90, loss: 0.0991290882229805
step: 100, loss: 0.036694541573524475
step: 110, loss: 0.02458631433546543
step: 120, loss: 0.06695941835641861
step: 130, loss: 0.041255079209804535
step: 140, loss: 0.0160845797508955
step: 150, loss: 0.022930379956960678
step: 160, loss: 0.06982962787151337
step: 170, loss: 0.0239556971937418
/usr/local/lib/python3.7/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=0.7425474254742549, f1=0.6596306068601583,
best_f1=0.6891566265060242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028772868681699038
step: 10, loss: 0.03168449550867081
step: 20, loss: 0.003129341173917055
step: 30, loss: 0.004621278028935194
step: 40, loss: 0.029182258993387222
step: 50, loss: 0.032347094267606735
step: 60, loss: 0.012271659448742867
step: 70, loss: 0.13849014043807983
step: 80, loss: 0.04809127002954483
step: 90, loss: 0.0009752122568897903
step: 100, loss: 0.14490115642547607
step: 110, loss: 0.0024730449076741934
step: 120, loss: 0.028240490704774857
step: 130, loss: 0.1767711043357849
step: 140, loss: 0.004547589924186468
step: 150, loss: 0.004451517015695572
step: 160, loss: 0.009483936242759228
step: 170, loss: 0.07319007068872452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7544303797468355, f1=0.6875, best_f1=0.6875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020692383870482445
step: 10, loss: 0.002484292956069112
step: 20, loss: 0.012767273001372814
step: 30, loss: 0.008640802465379238
step: 40, loss: 0.00026189247728325427
step: 50, loss: 0.0010154732735827565
step: 60, loss: 0.027266407385468483
step: 70, loss: 0.01956201158463955
step: 80, loss: 0.007243017200380564
step: 90, loss: 0.03141449764370918
step: 100, loss: 0.037373386323451996
step: 110, loss: 0.03237709775567055
step: 120, loss: 0.0017965127481147647
step: 130, loss: 0.005122466012835503
step: 140, loss: 0.04166979715228081
step: 150, loss: 0.011879796162247658
step: 160, loss: 0.013646155595779419
step: 170, loss: 0.0012797435047104955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7611548556430446, f1=0.6765432098765432,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015076100826263428
step: 10, loss: 0.00968840904533863
step: 20, loss: 0.0004940062644891441
step: 30, loss: 0.003560279496014118
step: 40, loss: 0.007098449394106865
step: 50, loss: 0.001211208407767117
step: 60, loss: 0.07017780095338821
step: 70, loss: 0.042450617998838425
step: 80, loss: 0.001767513225786388
step: 90, loss: 0.03181811794638634

```

```

step: 100, loss: 0.0032179132103919983
step: 110, loss: 0.001822007354348898
step: 120, loss: 0.12371838092803955
step: 130, loss: 0.03654395043849945
step: 140, loss: 0.0056238723918795586
step: 150, loss: 0.004871978424489498
step: 160, loss: 0.017549004405736923
step: 170, loss: 0.01583797298371792
/usr/local/lib/python3.7/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=0.7559055118110237, f1=0.6855670103092784,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03354783356189728
step: 10, loss: 0.0012065083719789982
step: 20, loss: 0.023009544238448143
step: 30, loss: 0.13776260614395142
step: 40, loss: 0.00031032165861688554
step: 50, loss: 0.001271998044103384
step: 60, loss: 0.001305223093368113
step: 70, loss: 0.028370968997478485
step: 80, loss: 0.00042955190292559564
step: 90, loss: 0.013544498011469841
step: 100, loss: 0.00028725009178742766
step: 110, loss: 0.028255203738808632
step: 120, loss: 0.001065578544512391
step: 130, loss: 0.002793444087728858
step: 140, loss: 0.07746372371912003
step: 150, loss: 0.024699252098798752
step: 160, loss: 0.036234524101018906
step: 170, loss: 0.006069280207157135
/usr/local/lib/python3.7/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=0.7349081364829397, f1=0.6865671641791046,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10612209886312485
step: 10, loss: 0.002861126558855176
step: 20, loss: 0.002123218961060047
step: 30, loss: 0.0006749752792529762
step: 40, loss: 0.0006674777250736952
step: 50, loss: 0.009155414067208767

```

```

step: 60, loss: 0.0965646505355835
step: 70, loss: 0.005555363371968269
step: 80, loss: 0.0004404261999297887
step: 90, loss: 0.0017991416389122605
step: 100, loss: 0.0005787370027974248
step: 110, loss: 0.15905094146728516
step: 120, loss: 0.0012548785889521241
step: 130, loss: 0.0023813049774616957
step: 140, loss: 0.003494460368528962
step: 150, loss: 0.006112012080848217
step: 160, loss: 0.02200733870267868
step: 170, loss: 0.0049991304986178875
/usr/local/lib/python3.7/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=0.7329842931937172, f1=0.6812652068126521,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004689149500336498
step: 10, loss: 0.0015856027603149414
step: 20, loss: 0.006759136449545622
step: 30, loss: 0.0010082870721817017
step: 40, loss: 0.0025459073949605227
step: 50, loss: 0.0003938714216928929
step: 60, loss: 0.019197987392544746
step: 70, loss: 0.022358929738402367
step: 80, loss: 0.00015368068125098944
step: 90, loss: 0.0006388541660271585
step: 100, loss: 0.0009108234080486
step: 110, loss: 0.006590315140783787
step: 120, loss: 0.04183662682771683
step: 130, loss: 0.0014758300967514515
step: 140, loss: 0.00262852874584496
step: 150, loss: 0.023599622771143913
step: 160, loss: 0.007709610275924206
step: 170, loss: 0.0006120817852206528
/usr/local/lib/python3.7/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=0.7238605898123324, f1=0.7022900763358779,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000818712986074388
step: 10, loss: 0.001346004195511341

```

```

step: 20, loss: 0.0029931417666375637
step: 30, loss: 0.00048106032772921026
step: 40, loss: 0.0007539608050137758
step: 50, loss: 0.008941726759076118
step: 60, loss: 0.0015178648754954338
step: 70, loss: 0.0014067908050492406
step: 80, loss: 0.001850320608355105
step: 90, loss: 0.00038418322219513357
step: 100, loss: 0.02925891801714897
step: 110, loss: 8.75144251040183e-05
step: 120, loss: 0.009767325595021248
step: 130, loss: 7.199410174507648e-05
step: 140, loss: 0.00038413709262385964
step: 150, loss: 0.0005043752025812864
step: 160, loss: 0.00028735995874740183
step: 170, loss: 0.002277128864079714
/usr/local/lib/python3.7/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=0.7282608695652175, f1=0.7049608355091384,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003902257012668997
step: 10, loss: 0.00010722235310822725
step: 20, loss: 0.0008086463203653693
step: 30, loss: 0.0009229448041878641
step: 40, loss: 0.0001419565815012902
step: 50, loss: 0.0001473233860451728
step: 60, loss: 0.001504651503637433
step: 70, loss: 0.0008912413031794131
step: 80, loss: 0.0013866299999615252
step: 90, loss: 0.0003998947504442185
step: 100, loss: 0.008808568120002747
step: 110, loss: 0.0002708864922169596
step: 120, loss: 0.006375140510499477
step: 130, loss: 0.019248586148023605
step: 140, loss: 0.0006029929500073195
step: 150, loss: 7.492581789847463e-05
step: 160, loss: 4.481776340981014e-05
step: 170, loss: 8.832142339088023e-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=0.7467362924281984, f1=0.7032418952618454,
best_f1=0.6765432098765432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024650455452501774
step: 10, loss: 0.010770321823656559
step: 20, loss: 0.0035095438361167908
step: 30, loss: 0.0006704897969029844
step: 40, loss: 0.0008904136484488845
step: 50, loss: 0.0058739641681313515
step: 60, loss: 0.010514365509152412
step: 70, loss: 6.555533036589622e-05
step: 80, loss: 0.001033502398058772
step: 90, loss: 0.0013039883924648166
step: 100, loss: 0.00043748709140345454
step: 110, loss: 0.00013066285464446992
step: 120, loss: 0.004694841336458921
step: 130, loss: 0.00041687514749355614
step: 140, loss: 0.00018318775983061641
step: 150, loss: 0.0006024149479344487
step: 160, loss: 0.00024240183120127767
step: 170, loss: 0.0004016159218735993
/usr/local/lib/python3.7/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=0.7454068241469816, f1=0.7043701799485862,
best_f1=0.6765432098765432
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:06, 291.17it/s]
load_f1 = 0.5204918032786885
real_f1 = 0.4932562620423892
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 274.55it/s]

```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

```

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

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

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Dirty/DBLP-ACM \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \
    --lm bert-base-uncased \

```

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

```
Downloading: 100% 433/433 [00:00<00:00, 376kB/s]  
Downloading: 100% 232k/232k [00:00<00:00, 253kB/s]  
Downloading: 100% 440M/440M [00:06<00:00, 64.0MB/s]  
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32  
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

```
step: 0, loss: 0.6056886315345764  
step: 10, loss: 0.6477164030075073  
step: 20, loss: 0.5119900703430176  
step: 30, loss: 0.5053315162658691
```



```

step: 40, loss: 0.34141796827316284
step: 50, loss: 0.11920282244682312
step: 60, loss: 0.06928717344999313
step: 70, loss: 0.05863545462489128
step: 80, loss: 0.1128811240196228
step: 90, loss: 0.08870967477560043
step: 100, loss: 0.00643910700455308
step: 110, loss: 0.12502220273017883
step: 120, loss: 0.03325754404067993
step: 130, loss: 0.010545316152274609
step: 140, loss: 0.028176141902804375
step: 150, loss: 0.03826598450541496
step: 160, loss: 0.007408057805150747
step: 170, loss: 0.25152692198753357
step: 180, loss: 0.0796140804886818
step: 190, loss: 0.06281603872776031
step: 200, loss: 0.06306150555610657
step: 210, loss: 0.019388064742088318
step: 220, loss: 0.014405859634280205
step: 230, loss: 0.06326166540384293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9508928571428572, f1=0.9592760180995475,
best_f1=0.9592760180995475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008791577070951462
step: 10, loss: 0.03250126913189888
step: 20, loss: 0.2226814329624176
step: 30, loss: 0.3092679977416992
step: 40, loss: 0.08200091123580933
step: 50, loss: 0.005605530459433794
step: 60, loss: 0.07906699180603027
step: 70, loss: 0.11032112687826157
step: 80, loss: 0.09134995192289352
step: 90, loss: 0.012800583615899086
step: 100, loss: 0.0578479990363121
step: 110, loss: 0.1459243893623352
step: 120, loss: 0.04309351369738579
step: 130, loss: 0.11459571123123169
step: 140, loss: 0.002224658615887165
step: 150, loss: 0.02547730877995491

```

```

step: 160, loss: 0.031063206493854523
step: 170, loss: 0.044052738696336746
step: 180, loss: 0.11072483658790588
step: 190, loss: 0.021165112033486366
step: 200, loss: 0.006745608989149332
step: 210, loss: 0.002472186926752329
step: 220, loss: 0.08070580661296844
step: 230, loss: 0.13104377686977386
/usr/local/lib/python3.7/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=0.9426048565121413, f1=0.931537598204265,
best_f1=0.9592760180995475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01691759191453457
step: 10, loss: 0.014123152941465378
step: 20, loss: 0.050486184656620026
step: 30, loss: 0.016101285815238953
step: 40, loss: 0.12799471616744995
step: 50, loss: 0.06502983719110489
step: 60, loss: 0.01078851893544197
step: 70, loss: 0.06711624562740326
step: 80, loss: 0.001951704267412424
step: 90, loss: 0.07047472149133682
step: 100, loss: 0.1631428599357605
step: 110, loss: 0.0017884352710098028
step: 120, loss: 0.008695611730217934
step: 130, loss: 0.001301920972764492
step: 140, loss: 0.030527396127581596
step: 150, loss: 0.046671267598867416
step: 160, loss: 0.024337556213140488
step: 170, loss: 0.03313437104225159
step: 180, loss: 0.061952855437994
step: 190, loss: 0.009832366369664669
step: 200, loss: 0.03783845156431198
step: 210, loss: 0.005356995388865471
step: 220, loss: 0.006451861932873726
step: 230, loss: 0.0007962350500747561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9581920903954803, f1=0.9476082004555808,

```

```

best_f1=0.9476082004555808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012707623653113842
step: 10, loss: 0.0009926398051902652
step: 20, loss: 0.0005195985431782901
step: 30, loss: 0.001970821525901556
step: 40, loss: 0.0028502659406512976
step: 50, loss: 0.0008518394315615296
step: 60, loss: 0.001441402593627572
step: 70, loss: 0.004948113113641739
step: 80, loss: 0.0014881190145388246
step: 90, loss: 0.014757232740521431
step: 100, loss: 0.0709032341837883
step: 110, loss: 0.0022736345417797565
step: 120, loss: 0.05791407451033592
step: 130, loss: 0.048039648681879044
step: 140, loss: 0.0032400607597082853
step: 150, loss: 0.281240314245224
step: 160, loss: 0.013970690779387951
step: 170, loss: 0.010973690077662468
step: 180, loss: 0.0013722728472203016
step: 190, loss: 0.012073108926415443
step: 200, loss: 0.0015954344999045134
step: 210, loss: 0.07271133363246918
step: 220, loss: 0.00218446203507483
step: 230, loss: 0.006318948231637478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9651293588301463, f1=0.9573033707865168,
best_f1=0.9573033707865168
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012565661454573274
step: 10, loss: 0.0031498614698648453
step: 20, loss: 0.006162989418953657
step: 30, loss: 0.0010558712529018521
step: 40, loss: 0.005690914113074541
step: 50, loss: 0.006712621543556452
step: 60, loss: 0.15412576496601105
step: 70, loss: 0.0009461325826123357

```

```

step: 80, loss: 0.003429712727665901
step: 90, loss: 0.03385109454393387
step: 100, loss: 0.0005808679852634668
step: 110, loss: 0.0006628865958191454
step: 120, loss: 0.0009869420900940895
step: 130, loss: 0.009230067953467369
step: 140, loss: 0.007560462225228548
step: 150, loss: 0.0018889947095885873
step: 160, loss: 0.0003822194703388959
step: 170, loss: 0.014951176010072231
step: 180, loss: 0.0019887047819793224
step: 190, loss: 0.1096482053399086
step: 200, loss: 0.010726679116487503
step: 210, loss: 0.005224545951932669
step: 220, loss: 0.042011458426713943
step: 230, loss: 0.0016551786102354527
/usr/local/lib/python3.7/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=0.9642857142857144, f1=0.9497206703910615,
best_f1=0.9573033707865168
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013059388846158981
step: 10, loss: 0.003180459374561906
step: 20, loss: 0.07019439339637756
step: 30, loss: 0.0003605192177928984
step: 40, loss: 0.0024054355453699827
step: 50, loss: 0.000447474216343835
step: 60, loss: 0.0004714185488410294
step: 70, loss: 0.0013633582275360823
step: 80, loss: 0.0011405693367123604
step: 90, loss: 0.0051711080595850945
step: 100, loss: 0.04996149614453316
step: 110, loss: 0.0008809567079879344
step: 120, loss: 0.001086438074707985
step: 130, loss: 0.0008517105598002672
step: 140, loss: 0.00120898790191859
step: 150, loss: 0.0052543580532073975
step: 160, loss: 0.016695650294423103
step: 170, loss: 0.00026127960882149637
step: 180, loss: 0.008161196485161781
step: 190, loss: 0.012583872303366661
step: 200, loss: 0.0008764928788878024
step: 210, loss: 0.00236326246522367
step: 220, loss: 0.0003632647276390344
step: 230, loss: 0.09339550882577896

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9655937846836848, f1=0.9507829977628636,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043829042464494705
step: 10, loss: 0.00019102550868410617
step: 20, loss: 0.00039476921665482223
step: 30, loss: 0.00031123351072892547
step: 40, loss: 0.0002912667114287615
step: 50, loss: 0.0004146495775785297
step: 60, loss: 0.004121480043977499
step: 70, loss: 0.01682838797569275
step: 80, loss: 0.013075508177280426
step: 90, loss: 0.00030519935535266995
step: 100, loss: 0.0007524718530476093
step: 110, loss: 0.013149912469089031
step: 120, loss: 0.00017930469766724855
step: 130, loss: 0.00018011065549217165
step: 140, loss: 0.0011173433158546686
step: 150, loss: 0.00034890740062110126
step: 160, loss: 0.09747380018234253
step: 170, loss: 0.0010547207202762365
step: 180, loss: 0.005928706377744675
step: 190, loss: 0.0002915715449489653
step: 200, loss: 0.07945563644170761
step: 210, loss: 0.00015720498049631715
step: 220, loss: 0.026302343234419823
step: 230, loss: 0.0010751902591437101
/usr/local/lib/python3.7/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=0.9608938547486034, f1=0.9516310461192351,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005383676616474986
step: 10, loss: 0.0016022762283682823
step: 20, loss: 0.008338849060237408
step: 30, loss: 0.0007017123280093074

```

```

step: 40, loss: 0.0037285704165697098
step: 50, loss: 0.000808928394690156
step: 60, loss: 0.000543877249583602
step: 70, loss: 0.0009467563359066844
step: 80, loss: 0.005893042776733637
step: 90, loss: 0.00330904102884233
step: 100, loss: 0.0030787871219217777
step: 110, loss: 0.0021122093312442303
step: 120, loss: 0.0006872161757200956
step: 130, loss: 0.0014307152014225721
step: 140, loss: 0.00032852301956154406
step: 150, loss: 0.0008712007547728717
step: 160, loss: 0.0011596172116696835
step: 170, loss: 0.0008585843606851995
step: 180, loss: 0.0015370770124718547
step: 190, loss: 0.005286881234496832
step: 200, loss: 0.002016456564888358
step: 210, loss: 0.005433120299130678
step: 220, loss: 0.001388011034578085
step: 230, loss: 0.001915643224492669
/usr/local/lib/python3.7/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=0.9579067121729239, f1=0.9465301478953355,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013125576078891754
step: 10, loss: 0.0022570763248950243
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.12611795961856842
step: 30, loss: 0.0001017262038658373
step: 40, loss: 0.09906216710805893
step: 50, loss: 0.0001905754761537537
step: 60, loss: 0.0009637639159336686
step: 70, loss: 0.00012030071957269683
step: 80, loss: 0.0009435968822799623
step: 90, loss: 0.0004346877394709736
step: 100, loss: 0.0009612567955628037
step: 110, loss: 0.00018495690892450511
step: 120, loss: 0.0005984918680042028
step: 130, loss: 0.00018528325017541647
step: 140, loss: 0.002005132846534252
step: 150, loss: 0.00045001041144132614
step: 160, loss: 0.0003233767638448626
step: 170, loss: 0.004722686484456062
step: 180, loss: 0.010366939939558506

```

```

step: 190, loss: 0.00022880153846926987
step: 200, loss: 0.00041190857882611454
step: 210, loss: 0.0001778937003109604
step: 220, loss: 0.00017195515101775527
step: 230, loss: 0.0011826584814116359
/usr/local/lib/python3.7/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=0.9592760180995475, f1=0.9462857142857144,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016895937733352184
step: 10, loss: 6.774830399081111e-05
step: 20, loss: 0.0002072489442070946
step: 30, loss: 0.00018164831271860749
step: 40, loss: 0.000221898706513457
step: 50, loss: 0.001403805916197598
step: 60, loss: 0.0003839356650132686
step: 70, loss: 0.0001507286069681868
step: 80, loss: 0.00013572101306635886
step: 90, loss: 0.0004344775516074151
step: 100, loss: 0.0012167464010417461
step: 110, loss: 0.0001861476484918967
step: 120, loss: 0.0021654332522302866
step: 130, loss: 0.00038520805537700653
step: 140, loss: 0.02220771089196205
step: 150, loss: 0.03626595437526703
step: 160, loss: 0.00011996504326816648
step: 170, loss: 9.906773630063981e-05
step: 180, loss: 0.0003515086427796632
step: 190, loss: 0.0013048367109149694
step: 200, loss: 0.0002109608321916312
step: 210, loss: 8.885315037332475e-05
step: 220, loss: 0.00010373729310231283
step: 230, loss: 0.0001487514964537695
/usr/local/lib/python3.7/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=0.9551569506726457, f1=0.9470124013528749,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008997551631182432
step: 10, loss: 0.00035926737473346293
step: 20, loss: 0.0005076415254734457

```

```

step: 30, loss: 0.057978205382823944
step: 40, loss: 0.011564713902771473
step: 50, loss: 0.00022409451776184142
step: 60, loss: 0.0005134175298735499
step: 70, loss: 0.0010936959879472852
step: 80, loss: 0.00015954130503814667
step: 90, loss: 0.007041020784527063
step: 100, loss: 0.0027490719221532345
step: 110, loss: 0.0004403524217195809
step: 120, loss: 0.0003657430352177471
step: 130, loss: 0.0013385799247771502
step: 140, loss: 0.0015335968928411603
step: 150, loss: 0.009333007968962193
step: 160, loss: 0.00010031298006651923
step: 170, loss: 0.040801092982292175
step: 180, loss: 0.0005783412489108741
step: 190, loss: 0.00010318938439013436
step: 200, loss: 0.000633508141618222
step: 210, loss: 0.00022574169270228595
step: 220, loss: 0.00017848171410150826
step: 230, loss: 0.0001868109975475818
/usr/local/lib/python3.7/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=0.9569060773480662, f1=0.9494505494505495,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012360165128484368
step: 10, loss: 0.0008704523788765073
step: 20, loss: 0.000509253703057766
step: 30, loss: 0.03255153074860573
step: 40, loss: 0.0014362073270604014
step: 50, loss: 0.0016826100181788206
step: 60, loss: 0.008240296505391598
step: 70, loss: 6.524587661260739e-05
step: 80, loss: 0.0004216005327180028
step: 90, loss: 0.00017924417625181377
step: 100, loss: 9.493820834904909e-05
step: 110, loss: 0.0001432339195162058
step: 120, loss: 0.0001899939961731434
step: 130, loss: 0.0028572892770171165
step: 140, loss: 0.00013773306272923946
step: 150, loss: 7.927157275844365e-05
step: 160, loss: 0.0003539232711773366
step: 170, loss: 0.0006881842855364084
step: 180, loss: 0.0012693193275481462

```



```

step: 190, loss: 0.00041675209649838507
step: 200, loss: 0.00010213711357209831
step: 210, loss: 0.00024406146258115768
step: 220, loss: 0.0002876035578083247
step: 230, loss: 0.006777593400329351
/usr/local/lib/python3.7/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=0.9581920903954803, f1=0.9479638009049774,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0052519566379487514
step: 10, loss: 9.420516289537773e-05
step: 20, loss: 0.0006558746099472046
step: 30, loss: 0.00038855941966176033
step: 40, loss: 0.00010164744162466377
step: 50, loss: 0.00014517392264679074
step: 60, loss: 0.00012402029824443161
step: 70, loss: 6.030278018442914e-05
step: 80, loss: 6.819751433795318e-05
step: 90, loss: 0.005664260592311621
step: 100, loss: 7.242141145979986e-05
step: 110, loss: 0.0034507049713283777
step: 120, loss: 0.09027349948883057
step: 130, loss: 0.05925688147544861
step: 140, loss: 0.00010787863720906898
step: 150, loss: 0.00011929919855901971
step: 160, loss: 0.00024531225790269673
step: 170, loss: 7.670898048672825e-05
step: 180, loss: 0.00013916095485910773
step: 190, loss: 5.112447252031416e-05
step: 200, loss: 7.823151827324182e-05
step: 210, loss: 3.820009442279115e-05
step: 220, loss: 6.165131344459951e-05
step: 230, loss: 5.8166471717413515e-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=0.9574944071588367, f1=0.9503386004514672,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.444190072827041e-05
step: 10, loss: 0.00014962385466787964
step: 20, loss: 9.725490963319317e-05

```

```

step: 30, loss: 8.537832763977349e-05
step: 40, loss: 0.009719189256429672
step: 50, loss: 0.00029854566673748195
step: 60, loss: 0.00011826503032352775
step: 70, loss: 0.00019748690829146653
step: 80, loss: 0.00013454060535877943
step: 90, loss: 0.0010749166831374168
step: 100, loss: 0.00020366060198284686
step: 110, loss: 0.00039010224281810224
step: 120, loss: 2.5275150619563647e-05
step: 130, loss: 0.00016836417489685118
step: 140, loss: 0.00010208610910922289
step: 150, loss: 5.426335701486096e-05
step: 160, loss: 0.001442533452063799
step: 170, loss: 2.34388280659914e-05
step: 180, loss: 7.650200132047758e-05
step: 190, loss: 0.00010371936514275149
step: 200, loss: 0.00010249627666780725
step: 210, loss: 0.000263432040810585
step: 220, loss: 0.00015921200974844396
step: 230, loss: 0.00013084286183584481
/usr/local/lib/python3.7/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=0.9597315436241611, f1=0.9496080627099663,
best_f1=0.9507829977628636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.941076647606678e-05
step: 10, loss: 4.596187500283122e-05
step: 20, loss: 0.00020746963855344802
step: 30, loss: 5.0708524213405326e-05
step: 40, loss: 9.537526784697548e-05
step: 50, loss: 0.026563363149762154
step: 60, loss: 0.0003934327978640795
step: 70, loss: 0.0001571676111780107
step: 80, loss: 0.0002408359869150445
step: 90, loss: 0.0001638206304050982
step: 100, loss: 7.132275641197339e-05
step: 110, loss: 6.078464866732247e-05
step: 120, loss: 0.00011523339344421402
step: 130, loss: 0.0003770582261495292
step: 140, loss: 5.2371597121236846e-05
step: 150, loss: 0.00012159966718172655
step: 160, loss: 4.8025529395090416e-05
step: 170, loss: 7.009552064118907e-05
step: 180, loss: 0.0005350380088202655

```

```

step: 190, loss: 0.0009472128585912287
step: 200, loss: 0.0004242589930072427
step: 210, loss: 6.420948193408549e-05
step: 220, loss: 0.0006134888390079141
step: 230, loss: 4.225711381877725e-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=0.9599109131403119, f1=0.9511111111111112,
best_f1=0.9507829977628636
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:15, 158.53it/s]
load_f1 = 0.9610678531701891
real_f1 = 0.9599109131403119
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 179.79it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6458385586738586
step: 10, loss: 0.5338491797447205
step: 20, loss: 0.5805912017822266
step: 30, loss: 0.32320523262023926
step: 40, loss: 0.2911936938762665
step: 50, loss: 0.1420029103755951
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.23036623001098633
step: 70, loss: 0.13080285489559174
step: 80, loss: 0.20285218954086304
step: 90, loss: 0.3509712219238281
step: 100, loss: 0.24363499879837036
step: 110, loss: 0.08518707752227783
step: 120, loss: 0.13717660307884216
step: 130, loss: 0.2524554133415222
step: 140, loss: 0.11530061811208725
step: 150, loss: 0.12420317530632019
step: 160, loss: 0.11213447898626328
step: 170, loss: 0.2630195617675781
step: 180, loss: 0.03360070288181305
step: 190, loss: 0.012237400747835636
step: 200, loss: 0.15835241973400116
step: 210, loss: 0.0640980526804924
step: 220, loss: 0.11617438495159149
step: 230, loss: 0.15422998368740082
step: 240, loss: 0.13924555480480194

```

```

step: 250, loss: 0.0680382028222084
step: 260, loss: 0.030272820964455605
step: 270, loss: 0.11121238768100739
step: 280, loss: 0.13781870901584625
step: 290, loss: 0.17780816555023193
step: 300, loss: 0.06127922236919403
step: 310, loss: 0.14937879145145416
step: 320, loss: 0.09536803513765335
step: 330, loss: 0.108514204621315
step: 340, loss: 0.13143853843212128
step: 350, loss: 0.06371571868658066
step: 360, loss: 0.1413014531135559
step: 370, loss: 0.13004370033740997
step: 380, loss: 0.02654033713042736
step: 390, loss: 0.22187069058418274
step: 400, loss: 0.26112645864486694
step: 410, loss: 0.07726050168275833
step: 420, loss: 0.10355448722839355
step: 430, loss: 0.1541513055562973
step: 440, loss: 0.09551849216222763
step: 450, loss: 0.014758805744349957
step: 460, loss: 0.04987778887152672
step: 470, loss: 0.0916495993733406
step: 480, loss: 0.10287825018167496
step: 490, loss: 0.13377881050109863
step: 500, loss: 0.07660434395074844
step: 510, loss: 0.09184859693050385
step: 520, loss: 0.24766786396503448
step: 530, loss: 0.013968085870146751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.906791569086651, f1=0.8881578947368421,
best_f1=0.8881578947368421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12449858337640762
step: 10, loss: 0.1057661920785904
step: 20, loss: 0.07745780795812607
step: 30, loss: 0.08267044275999069
step: 40, loss: 0.09026886522769928
step: 50, loss: 0.14782637357711792
step: 60, loss: 0.022903235629200935

```

step: 70, loss: 0.035274017602205276
step: 80, loss: 0.10413920134305954
step: 90, loss: 0.09359388053417206
step: 100, loss: 0.10704532265663147
step: 110, loss: 0.07330134510993958
step: 120, loss: 0.04367583245038986
step: 130, loss: 0.13500146567821503
step: 140, loss: 0.08358009159564972
step: 150, loss: 0.15675979852676392
step: 160, loss: 0.029556291177868843
step: 170, loss: 0.19172883033752441
step: 180, loss: 0.02368447557091713
step: 190, loss: 0.10886456072330475
step: 200, loss: 0.03720122575759888
step: 210, loss: 0.11750950664281845
step: 220, loss: 0.057887837290763855
step: 230, loss: 0.07031211256980896
step: 240, loss: 0.09981245547533035
step: 250, loss: 0.08514942973852158
step: 260, loss: 0.029104411602020264
step: 270, loss: 0.2700631618499756
step: 280, loss: 0.16671490669250488
step: 290, loss: 0.02721807174384594
step: 300, loss: 0.18244080245494843
step: 310, loss: 0.022900035604834557
step: 320, loss: 0.17586617171764374
step: 330, loss: 0.09788604080677032
step: 340, loss: 0.050743404775857925
step: 350, loss: 0.005403914954513311
step: 360, loss: 0.19457566738128662
step: 370, loss: 0.27179044485092163
step: 380, loss: 0.1398356705904007
step: 390, loss: 0.15574456751346588
step: 400, loss: 0.028127169236540794
step: 410, loss: 0.09113152325153351
step: 420, loss: 0.04497024416923523
step: 430, loss: 0.20488138496875763
step: 440, loss: 0.08994040638208389
step: 450, loss: 0.034059759229421616
step: 460, loss: 0.13068531453609467
step: 470, loss: 0.07804366201162338
step: 480, loss: 0.16524578630924225
step: 490, loss: 0.047712989151477814
step: 500, loss: 0.2867357134819031
step: 510, loss: 0.028550703078508377
step: 520, loss: 0.13490645587444305
step: 530, loss: 0.04069199413061142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9133211678832116, f1=0.9132629646626893,
best_f1=0.9132629646626893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09132565557956696
step: 10, loss: 0.07833033055067062
step: 20, loss: 0.16270138323307037
step: 30, loss: 0.30612432956695557
step: 40, loss: 0.03282702714204788
step: 50, loss: 0.12963137030601501
step: 60, loss: 0.02937251515686512
step: 70, loss: 0.05234447494149208
step: 80, loss: 0.020953986793756485
step: 90, loss: 0.03627201169729233
step: 100, loss: 0.03244626149535179
step: 110, loss: 0.0056075481697916985
step: 120, loss: 0.044847991317510605
step: 130, loss: 0.005671561695635319
step: 140, loss: 0.008839139714837074
step: 150, loss: 0.1020573154091835
step: 160, loss: 0.04377405345439911
step: 170, loss: 0.09695538878440857
step: 180, loss: 0.06667955219745636
step: 190, loss: 0.1380840390920639
step: 200, loss: 0.02047593891620636
step: 210, loss: 0.051554542034864426
step: 220, loss: 0.09937074035406113
step: 230, loss: 0.21525363624095917
step: 240, loss: 0.1022290512919426
step: 250, loss: 0.007971487008035183
step: 260, loss: 0.047472160309553146
step: 270, loss: 0.07286716252565384
step: 280, loss: 0.16142994165420532
step: 290, loss: 0.0134092066437006
step: 300, loss: 0.07934559881687164
step: 310, loss: 0.1464604288339615
step: 320, loss: 0.05847702920436859
step: 330, loss: 0.012587637640535831
step: 340, loss: 0.014835579320788383
step: 350, loss: 0.02242213673889637
step: 360, loss: 0.06284986436367035

```



```

step: 370, loss: 0.029063493013381958
step: 380, loss: 0.008305469527840614
step: 390, loss: 0.06261922419071198
step: 400, loss: 0.043435513973236084
step: 410, loss: 0.009080016054213047
step: 420, loss: 0.12573584914207458
step: 430, loss: 0.008249152451753616
step: 440, loss: 0.10962120443582535
step: 450, loss: 0.04847179353237152
step: 460, loss: 0.10932448506355286
step: 470, loss: 0.04171104356646538
step: 480, loss: 0.15129530429840088
step: 490, loss: 0.013435701839625835
step: 500, loss: 0.14163537323474884
step: 510, loss: 0.016634037718176842
step: 520, loss: 0.14002768695354462
step: 530, loss: 0.20981276035308838
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.914577530176416, f1=0.9064409967089797,
best_f1=0.9064409967089797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04798051714897156
step: 10, loss: 0.014451922848820686
step: 20, loss: 0.0037016139831393957
step: 30, loss: 0.059354811906814575
step: 40, loss: 0.05320942774415016
step: 50, loss: 0.07061005383729935
step: 60, loss: 0.10496469587087631
step: 70, loss: 0.050061389803886414
step: 80, loss: 0.006152356509119272
step: 90, loss: 0.14257945120334625
step: 100, loss: 0.0018942009191960096
step: 110, loss: 0.2046266496181488
step: 120, loss: 0.0017642317106947303
step: 130, loss: 0.0016201342223212123
step: 140, loss: 0.03022405132651329
step: 150, loss: 0.06149066239595413
step: 160, loss: 0.03461673855781555
step: 170, loss: 0.013833894394338131
step: 180, loss: 0.028697310015559196

```

```

step: 190, loss: 0.04491313919425011
step: 200, loss: 0.028800198808312416
step: 210, loss: 0.04692227765917778
step: 220, loss: 0.0065596760250627995
step: 230, loss: 0.044121332466602325
step: 240, loss: 0.007828169502317905
step: 250, loss: 0.11717624962329865
step: 260, loss: 0.006144512910395861
step: 270, loss: 0.04896538704633713
step: 280, loss: 0.16461993753910065
step: 290, loss: 0.02226891927421093
step: 300, loss: 0.005601199809461832
step: 310, loss: 0.007879525423049927
step: 320, loss: 0.004625466652214527
step: 330, loss: 0.027457047253847122
step: 340, loss: 0.13294605910778046
step: 350, loss: 0.039502374827861786
step: 360, loss: 0.08144433796405792
step: 370, loss: 0.008571477606892586
step: 380, loss: 0.003808780573308468
step: 390, loss: 0.006867063231766224
step: 400, loss: 0.018401991575956345
step: 410, loss: 0.00791771337389946
step: 420, loss: 0.010524297133088112
step: 430, loss: 0.05550113692879677
step: 440, loss: 0.006567536387592554
step: 450, loss: 0.032040953636169434
step: 460, loss: 0.011032812297344208
step: 470, loss: 0.034533146768808365
step: 480, loss: 0.08920101821422577
step: 490, loss: 0.0013748033670708537
step: 500, loss: 0.019132595509290695
step: 510, loss: 0.05825566500425339
step: 520, loss: 0.03849392756819725
step: 530, loss: 0.03156854957342148
/usr/local/lib/python3.7/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=0.9124767225325885, f1=0.9059429106223678,
best_f1=0.9064409967089797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029038092121481895
step: 10, loss: 0.06101314350962639
step: 20, loss: 0.0025150151923298836
step: 30, loss: 0.002837030915543437
step: 40, loss: 0.02065243199467659

```

step: 50, loss: 0.058233052492141724
step: 60, loss: 0.09751803427934647
step: 70, loss: 0.045642051845788956
step: 80, loss: 0.002021005842834711
step: 90, loss: 0.005949602462351322
step: 100, loss: 0.007328774780035019
step: 110, loss: 0.0011364702368155122
step: 120, loss: 0.0012851146748289466
step: 130, loss: 0.002462235512211919
step: 140, loss: 0.002770782681182027
step: 150, loss: 0.01205619890242815
step: 160, loss: 0.0008937412058003247
step: 170, loss: 0.009866036474704742
step: 180, loss: 0.002718702657148242
step: 190, loss: 0.002024558838456869
step: 200, loss: 0.008248649537563324
step: 210, loss: 0.19636496901512146
step: 220, loss: 0.0849122628569603
step: 230, loss: 0.05562504380941391
step: 240, loss: 0.0086777089163661
step: 250, loss: 0.07553191483020782
step: 260, loss: 0.0011847586138173938
step: 270, loss: 0.0011515025980770588
step: 280, loss: 0.010508504696190357
step: 290, loss: 0.17371374368667603
step: 300, loss: 0.008258906193077564
step: 310, loss: 0.004787803627550602
step: 320, loss: 0.16721034049987793
step: 330, loss: 0.03288843482732773
step: 340, loss: 0.006091021001338959
step: 350, loss: 0.005629411898553371
step: 360, loss: 0.01726541481912136
step: 370, loss: 0.014443854801356792
step: 380, loss: 0.0044230083003640175
step: 390, loss: 0.0014424691908061504
step: 400, loss: 0.1290895938873291
step: 410, loss: 0.020754659548401833
step: 420, loss: 0.0023669307120144367
step: 430, loss: 0.08661311864852905
step: 440, loss: 0.22665369510650635
step: 450, loss: 0.026165412738919258
step: 460, loss: 0.06595437228679657
step: 470, loss: 0.002959584351629019
step: 480, loss: 0.007976804859936237
step: 490, loss: 0.0017417063936591148
step: 500, loss: 0.024164168164134026
step: 510, loss: 0.0472676120698452
step: 520, loss: 0.0012829243205487728

```

step: 530, loss: 0.1616525501012802
/usr/local/lib/python3.7/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=0.9080622347949081, f1=0.9013685700802266,
best_f1=0.9064409967089797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006041083950549364
step: 10, loss: 0.004495135974138975
step: 20, loss: 0.12814758718013763
step: 30, loss: 0.016843143850564957
step: 40, loss: 0.057601917535066605
step: 50, loss: 0.0017118655377998948
step: 60, loss: 0.014007806777954102
step: 70, loss: 0.005854525603353977
step: 80, loss: 0.08436269313097
step: 90, loss: 0.004862921312451363
step: 100, loss: 0.0011310018599033356
step: 110, loss: 0.012077059596776962
step: 120, loss: 0.05601564794778824
step: 130, loss: 0.006117447279393673
step: 140, loss: 0.0029290567617863417
step: 150, loss: 0.0038010869175195694
step: 160, loss: 0.004460807424038649
step: 170, loss: 0.011853020638227463
step: 180, loss: 0.014020736329257488
step: 190, loss: 0.05828911438584328
step: 200, loss: 0.0022132531739771366
step: 210, loss: 0.012520567514002323
step: 220, loss: 0.0010864256182685494
step: 230, loss: 0.00374305690638721
step: 240, loss: 0.009433148428797722
step: 250, loss: 0.01307608000934124
step: 260, loss: 0.02349051646888256
step: 270, loss: 0.025757353752851486
step: 280, loss: 0.002936879638582468
step: 290, loss: 0.0005743193905800581
step: 300, loss: 0.0006558529566973448
step: 310, loss: 0.015501422807574272
step: 320, loss: 0.002446679165586829
step: 330, loss: 0.06813430786132812
step: 340, loss: 0.09142213314771652
step: 350, loss: 0.04491430148482323
step: 360, loss: 0.011356320232152939
step: 370, loss: 0.006584004499018192
step: 380, loss: 0.0006219608476385474

```

```

step: 390, loss: 0.02113204076886177
step: 400, loss: 0.029227830469608307
step: 410, loss: 0.004209303297102451
step: 420, loss: 0.08887719362974167
step: 430, loss: 0.0024699026253074408
step: 440, loss: 0.00033088645432144403
step: 450, loss: 0.0018950414378196
step: 460, loss: 0.002837776206433773
step: 470, loss: 0.004716121591627598
step: 480, loss: 0.015370012260973454
step: 490, loss: 0.009916224516928196
step: 500, loss: 0.015606631524860859
step: 510, loss: 0.07715202122926712
step: 520, loss: 0.05319838970899582
step: 530, loss: 0.004160596057772636
/usr/local/lib/python3.7/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=0.9006993006993006, f1=0.8928901200369345,
best_f1=0.9064409967089797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004805175121873617
step: 10, loss: 0.0013121768133714795
step: 20, loss: 0.008428615517914295
step: 30, loss: 0.04876039922237396
step: 40, loss: 0.002440560609102249
step: 50, loss: 0.1952071189880371
step: 60, loss: 0.023958489298820496
step: 70, loss: 0.009345454163849354
step: 80, loss: 0.007519339676946402
step: 90, loss: 0.020557915791869164
step: 100, loss: 0.0004443157813511789
step: 110, loss: 0.003402275964617729
step: 120, loss: 0.049967944622039795
step: 130, loss: 0.08997242152690887
step: 140, loss: 0.005978003144264221
step: 150, loss: 0.0009402782307006419
step: 160, loss: 0.0005894734640605748
step: 170, loss: 0.0033343906980007887
step: 180, loss: 0.00031587283592671156
step: 190, loss: 0.0017458497313782573
step: 200, loss: 0.0031008499208837748
step: 210, loss: 0.008258079178631306
step: 220, loss: 0.0010296928230673075
step: 230, loss: 0.052745599299669266
step: 240, loss: 0.028642408549785614

```

```

step: 250, loss: 0.00344344531185925
step: 260, loss: 0.0008169428328983486
step: 270, loss: 0.00033440583501942456
step: 280, loss: 0.0030733332969248295
step: 290, loss: 0.004029001109302044
step: 300, loss: 0.00473259761929512
step: 310, loss: 0.003790108487010002
step: 320, loss: 0.0016910959966480732
step: 330, loss: 0.017002491280436516
step: 340, loss: 0.04070466756820679
step: 350, loss: 0.01998104713857174
step: 360, loss: 0.0015512867830693722
step: 370, loss: 0.0014048166340216994
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 380, loss: 0.0031630361918359995
step: 390, loss: 0.00032040252699516714
step: 400, loss: 0.017226388677954674
step: 410, loss: 0.019377872347831726
step: 420, loss: 0.0051532406359910965
step: 430, loss: 0.005028167739510536
step: 440, loss: 0.035996437072753906
step: 450, loss: 0.0015243447851389647
step: 460, loss: 0.00039459875551983714
step: 470, loss: 0.016120711341500282
step: 480, loss: 0.06319751590490341
step: 490, loss: 0.017772266641259193
step: 500, loss: 0.0024981407914310694
step: 510, loss: 0.04945256561040878
step: 520, loss: 0.02072226069867611
step: 530, loss: 0.04202684387564659
/usr/local/lib/python3.7/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=0.9138090824837812, f1=0.9085027726432533,
best_f1=0.9064409967089797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023090558126568794
step: 10, loss: 0.006934137549251318
step: 20, loss: 0.00414546113461256
step: 30, loss: 0.013105612248182297
step: 40, loss: 0.0019846856594085693
step: 50, loss: 0.01465801801532507
step: 60, loss: 0.0026403784286230803
step: 70, loss: 0.005605196580290794
step: 80, loss: 0.003953332081437111
step: 90, loss: 0.0011110022896900773

```

```
step: 100, loss: 0.009832869283854961
step: 110, loss: 0.003015324240550399
step: 120, loss: 0.0052505554631352425
step: 130, loss: 0.0003380505077075213
step: 140, loss: 0.08826810121536255
step: 150, loss: 0.0033203568309545517
step: 160, loss: 0.012080434709787369
step: 170, loss: 0.11149918287992477
step: 180, loss: 0.00497865118086338
step: 190, loss: 0.0037183475214987993
step: 200, loss: 0.0024744493421167135
step: 210, loss: 0.018024714663624763
step: 220, loss: 0.0027993100229650736
step: 230, loss: 0.006868044380098581
step: 240, loss: 0.0016383016481995583
step: 250, loss: 0.009339312091469765
step: 260, loss: 0.0007106643752194941
step: 270, loss: 0.0015827217139303684
step: 280, loss: 0.009982790797948837
step: 290, loss: 0.0003836791729554534
step: 300, loss: 0.040052466094493866
step: 310, loss: 0.023279013112187386
step: 320, loss: 0.0013187223812565207
step: 330, loss: 0.0027119156438857317
step: 340, loss: 0.028465943410992622
step: 350, loss: 0.0022646707948297262
step: 360, loss: 0.009953036904335022
step: 370, loss: 0.0007646558806300163
step: 380, loss: 0.0025919212494045496
step: 390, loss: 0.0012604141375049949
step: 400, loss: 0.0012521882308647037
step: 410, loss: 0.0002636425488162786
step: 420, loss: 0.016754431650042534
step: 430, loss: 0.0317801758646965
step: 440, loss: 0.02518703043460846
step: 450, loss: 0.0029770899564027786
step: 460, loss: 0.0006530483369715512
step: 470, loss: 0.0005849734297953546
step: 480, loss: 0.0001486189430579543
step: 490, loss: 0.04242865741252899
step: 500, loss: 0.06720563769340515
step: 510, loss: 0.007988535799086094
step: 520, loss: 0.010032502934336662
step: 530, loss: 0.0006304431008175015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9155389640690621, f1=0.9034965034965035,

best_f1=0.9034965034965035

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

An input tensor was not cuda.

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

step: 0, loss: 0.0019205452408641577

step: 10, loss: 0.0020321847405284643

step: 20, loss: 0.010260287672281265

step: 30, loss: 0.01562872715294361

step: 40, loss: 0.0015924533363431692

step: 50, loss: 0.0016583132091909647

step: 60, loss: 0.0005927662132307887

step: 70, loss: 0.004344901069998741

step: 80, loss: 0.06516478955745697

step: 90, loss: 5.847239663125947e-05

step: 100, loss: 0.000995595008134842

step: 110, loss: 0.00038328184746205807

step: 120, loss: 0.018082158640027046

step: 130, loss: 0.007334609981626272

step: 140, loss: 0.0005316672031767666

step: 150, loss: 0.005117754451930523

step: 160, loss: 0.00045882147969678044

step: 170, loss: 0.05364629998803139

step: 180, loss: 0.0024617535527795553

step: 190, loss: 0.007009307388216257

step: 200, loss: 0.0017369546694681048

step: 210, loss: 0.001283105812035501

step: 220, loss: 0.005963183008134365

step: 230, loss: 0.0025895913131535053

step: 240, loss: 0.00023526183213107288

step: 250, loss: 0.0007805557106621563

step: 260, loss: 0.022066479548811913

step: 270, loss: 0.0002646348439157009

step: 280, loss: 0.004439723677933216

step: 290, loss: 0.054826125502586365

step: 300, loss: 0.0006471697124652565

step: 310, loss: 0.014956002123653889

step: 320, loss: 0.004209577571600676

step: 330, loss: 0.0009889568900689483

step: 340, loss: 0.0018904133467003703

step: 350, loss: 0.0009344038553535938

step: 360, loss: 8.798010821919888e-05

step: 370, loss: 0.001963237766176462

step: 380, loss: 0.0005121475551277399

step: 390, loss: 0.00040684948908165097


```

step: 400, loss: 0.17187553644180298
step: 410, loss: 0.034193072468042374
step: 420, loss: 0.013735871762037277
step: 430, loss: 0.006574103608727455
step: 440, loss: 0.0039932397194206715
step: 450, loss: 0.0071909986436367035
step: 460, loss: 0.0005043564597144723
step: 470, loss: 0.00022647310106549412
step: 480, loss: 0.00013925525126978755
step: 490, loss: 0.010424632579088211
step: 500, loss: 0.26576241850852966
step: 510, loss: 0.1657240390777588
step: 520, loss: 0.11350519955158234
step: 530, loss: 0.012460428290069103
/usr/local/lib/python3.7/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=0.9147647880763857, f1=0.9067164179104478,
best_f1=0.9034965034965035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008332473807968199
step: 10, loss: 0.000949582492467016
step: 20, loss: 0.00017324801592621952
step: 30, loss: 0.00023791154671926051
step: 40, loss: 0.00027849647449329495
step: 50, loss: 0.006483228411525488
step: 60, loss: 0.00035917063360102475
step: 70, loss: 0.0002398401265963912
step: 80, loss: 9.695608605397865e-05
step: 90, loss: 0.00027495561516843736
step: 100, loss: 6.662191299255937e-05
step: 110, loss: 7.084740354912356e-05
step: 120, loss: 0.0024430672638118267
step: 130, loss: 0.0026684908661991358
step: 140, loss: 0.007404182571917772
step: 150, loss: 0.0003897899005096406
step: 160, loss: 0.0013074291637167335
step: 170, loss: 0.0007537979981862009
step: 180, loss: 0.03394049033522606
step: 190, loss: 0.01076553575694561
step: 200, loss: 0.0017264883499592543
step: 210, loss: 0.0021107876673340797
step: 220, loss: 0.0028354593086987734
step: 230, loss: 0.006908917799592018
step: 240, loss: 0.0002766929392237216
step: 250, loss: 6.590421980945393e-05

```

```

step: 260, loss: 0.00579454330727458
step: 270, loss: 0.017694950103759766
step: 280, loss: 0.0010285940952599049
step: 290, loss: 0.0006640144856646657
step: 300, loss: 3.3794967748690397e-05
step: 310, loss: 9.178930486086756e-05
step: 320, loss: 0.00032341352198272943
step: 330, loss: 0.0007276355754584074
step: 340, loss: 0.055937591940164566
step: 350, loss: 0.016179069876670837
step: 360, loss: 0.05812479183077812
step: 370, loss: 0.0009474153048358858
step: 380, loss: 0.02332283742725849
step: 390, loss: 0.0018668259726837277
step: 400, loss: 0.002711877692490816
step: 410, loss: 0.004941301885992289
step: 420, loss: 0.00019592768512666225
step: 430, loss: 9.603925718693063e-05
step: 440, loss: 0.011901958845555782
step: 450, loss: 0.0013095731846988201
step: 460, loss: 0.005328099709004164
step: 470, loss: 0.00022757667466066778
step: 480, loss: 0.031927645206451416
step: 490, loss: 0.003922203090041876
step: 500, loss: 0.04657768830657005
step: 510, loss: 0.0007117468630895019
step: 520, loss: 0.05608082562685013
step: 530, loss: 0.0002200812887167558
/usr/local/lib/python3.7/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=0.9102623101702715, f1=0.9018518518518519,
best_f1=0.9034965034965035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035832656431011856
step: 10, loss: 0.0008682691841386259
step: 20, loss: 0.023672016337513924
step: 30, loss: 0.0003403203736525029
step: 40, loss: 0.0007406921358779073
step: 50, loss: 0.024525988847017288
step: 60, loss: 0.00022608730068895966
step: 70, loss: 0.0010124711552634835
step: 80, loss: 0.0004696745309047401
step: 90, loss: 0.00010468398249940947
step: 100, loss: 0.000695700931828469
step: 110, loss: 0.002654790412634611

```

```
step: 120, loss: 0.00032251118682324886
step: 130, loss: 0.02314835600554943
step: 140, loss: 0.008625196292996407
step: 150, loss: 0.00038234301609918475
step: 160, loss: 0.002233789535239339
step: 170, loss: 0.0021081336308270693
step: 180, loss: 0.0007377789006568491
step: 190, loss: 0.0005823103711009026
step: 200, loss: 0.053749363869428635
step: 210, loss: 0.03662165254354477
step: 220, loss: 0.00037990909186191857
step: 230, loss: 0.0004265836614649743
step: 240, loss: 0.0004072236188221723
step: 250, loss: 0.00013269938062876463
step: 260, loss: 0.0004313043027650565
step: 270, loss: 0.0007613276829943061
step: 280, loss: 0.00036897964309901
step: 290, loss: 0.00017694069538265467
step: 300, loss: 0.00042575254337862134
step: 310, loss: 0.0006582794594578445
step: 320, loss: 0.002688029548153281
step: 330, loss: 0.0015739199006929994
step: 340, loss: 0.0030872002243995667
step: 350, loss: 0.0001273232337553054
step: 360, loss: 0.00032748345984146
step: 370, loss: 0.0039435564540326595
step: 380, loss: 0.00011361181532265618
step: 390, loss: 0.0006013691890984774
step: 400, loss: 0.00017383597150910646
step: 410, loss: 0.000131098015117459
step: 420, loss: 0.0005274260765872896
step: 430, loss: 0.00013351930829230696
step: 440, loss: 0.003929690923541784
step: 450, loss: 0.00021937601559329778
step: 460, loss: 0.010074689984321594
step: 470, loss: 7.007303065620363e-05
step: 480, loss: 0.009461279958486557
step: 490, loss: 0.000247608550125733
step: 500, loss: 0.008420647121965885
step: 510, loss: 0.0013716586399823427
step: 520, loss: 0.00011602242011576891
step: 530, loss: 0.00018200530030298978
/usr/local/lib/python3.7/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=0.9105312208760485, f1=0.8967772069126577,
best_f1=0.9034965034965035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 8.97460849955678e-05
step: 10, loss: 8.14693485153839e-05
step: 20, loss: 0.052774522453546524
step: 30, loss: 0.0032493053004145622
step: 40, loss: 0.0015448906924575567
step: 50, loss: 0.02280329540371895
step: 60, loss: 0.001460028812289238
step: 70, loss: 0.001903186785057187
step: 80, loss: 0.0024458379484713078
step: 90, loss: 0.0007741920999251306
step: 100, loss: 0.000828603224363178
step: 110, loss: 0.00036368417204357684
step: 120, loss: 0.00034234108170494437
step: 130, loss: 0.00021369129535742104
step: 140, loss: 0.06661307066679001
step: 150, loss: 0.00021022545115556568
step: 160, loss: 7.546530832769349e-05
step: 170, loss: 0.0001614147622603923
step: 180, loss: 0.0004933848977088928
step: 190, loss: 0.0010242339922115207
step: 200, loss: 0.00029796166927553713
step: 210, loss: 0.0002919115941040218
step: 220, loss: 0.00017078670498449355
step: 230, loss: 7.642123819096014e-05
step: 240, loss: 0.0010845898650586605
step: 250, loss: 0.0008216479327529669
step: 260, loss: 0.00018258762429468334
step: 270, loss: 0.003268919885158539
step: 280, loss: 0.006662317086011171
step: 290, loss: 0.004936544690281153
step: 300, loss: 0.043091997504234314
step: 310, loss: 0.0002650014066603035
step: 320, loss: 7.679367263335735e-05
step: 330, loss: 0.00021815770014654845
step: 340, loss: 0.0001304163597524166
step: 350, loss: 0.0010462601203471422
step: 360, loss: 0.005229260306805372
step: 370, loss: 0.00045377176138572395
step: 380, loss: 0.00014933405327610672
step: 390, loss: 0.06566817313432693
step: 400, loss: 0.00027821699040941894
step: 410, loss: 0.007096480578184128
step: 420, loss: 0.013733072206377983
step: 430, loss: 0.0036107178311794996
step: 440, loss: 0.04889993369579315
step: 450, loss: 0.00023194460663944483
```

```

step: 460, loss: 0.00033138005528599024
step: 470, loss: 3.678564462461509e-05
step: 480, loss: 0.002382115926593542
step: 490, loss: 0.0021021494176238775
step: 500, loss: 0.0009986197110265493
step: 510, loss: 0.0013678919058293104
step: 520, loss: 4.2214396671624854e-05
step: 530, loss: 0.015458393841981888
/usr/local/lib/python3.7/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=0.9145220588235294, f1=0.9015639374425022,
best_f1=0.9034965034965035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00078775075962767
step: 10, loss: 0.0026306549552828074
step: 20, loss: 0.00306111597456038
step: 30, loss: 0.0007773960824124515
step: 40, loss: 7.130208541639149e-05
step: 50, loss: 0.007943719625473022
step: 60, loss: 0.0005996141117066145
step: 70, loss: 0.00014908387674950063
step: 80, loss: 3.455837577348575e-05
step: 90, loss: 0.00020227496861480176
step: 100, loss: 0.0014098206302151084
step: 110, loss: 0.20001405477523804
step: 120, loss: 0.0345681793987751
step: 130, loss: 0.015327511355280876
step: 140, loss: 0.00014639433356933296
step: 150, loss: 0.0018822327256202698
step: 160, loss: 2.3770631742081605e-05
step: 170, loss: 0.00010543897951720282
step: 180, loss: 0.0001734977267915383
step: 190, loss: 0.001292289001867175
step: 200, loss: 0.00016623119881842285
step: 210, loss: 0.00014317623572424054
step: 220, loss: 0.0017074085772037506
step: 230, loss: 0.0004558066721074283
step: 240, loss: 0.0012097125872969627
step: 250, loss: 0.0001414008584106341
step: 260, loss: 0.00018848455511033535
step: 270, loss: 0.00010354626283515245
step: 280, loss: 0.0001956548949237913
step: 290, loss: 0.0031006946228444576
step: 300, loss: 0.00010127530549652874
step: 310, loss: 0.0010271266801282763

```

```

step: 320, loss: 0.02574244514107704
step: 330, loss: 0.00030423171119764447
step: 340, loss: 0.00024359379312954843
step: 350, loss: 0.00034441036405041814
step: 360, loss: 9.818997205002233e-05
step: 370, loss: 0.00031017803121358156
step: 380, loss: 0.0034730995539575815
step: 390, loss: 0.0030654179863631725
step: 400, loss: 0.0006015911931172013
step: 410, loss: 0.0005170984077267349
step: 420, loss: 9.571530245011672e-05
step: 430, loss: 0.14264874160289764
step: 440, loss: 0.0006662991363555193
step: 450, loss: 0.00028600639780052006
step: 460, loss: 0.008868738077580929
step: 470, loss: 0.004190518520772457
step: 480, loss: 0.0003587030805647373
step: 490, loss: 0.012986204586923122
step: 500, loss: 6.129020766820759e-05
step: 510, loss: 0.003878260962665081
step: 520, loss: 0.0048503936268389225
step: 530, loss: 4.4188629544805735e-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=0.9115582592419279, f1=0.896519285042333,
best_f1=0.9034965034965035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005876236828044057
step: 10, loss: 3.320588803035207e-05
step: 20, loss: 0.0007089218124747276
step: 30, loss: 4.2092920921277255e-05
step: 40, loss: 0.0006795040098950267
step: 50, loss: 0.07961481809616089
step: 60, loss: 7.480999920517206e-05
step: 70, loss: 0.0026466669514775276
step: 80, loss: 4.2083102016476914e-05
step: 90, loss: 0.00017675297567620873
step: 100, loss: 0.01120050810277462
step: 110, loss: 0.0001446077658329159
step: 120, loss: 0.004434784408658743
step: 130, loss: 0.00191485695540905
step: 140, loss: 0.00011395322508178651
step: 150, loss: 6.019444117555395e-05
step: 160, loss: 3.8293721445370466e-05
step: 170, loss: 9.981430048355833e-05

```

```

step: 180, loss: 0.0012019608402624726
step: 190, loss: 0.00017635671247262508
step: 200, loss: 0.00013329244393389672
step: 210, loss: 0.0005249240784905851
step: 220, loss: 0.0005575857940129936
step: 230, loss: 0.028785547241568565
step: 240, loss: 0.00012455822434276342
step: 250, loss: 0.0029388079419732094
step: 260, loss: 0.0004998167860321701
step: 270, loss: 0.0019084591185674071
step: 280, loss: 0.000456327194115147
step: 290, loss: 0.0020066658034920692
step: 300, loss: 0.00023746326041873544
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 310, loss: 0.04258139431476593
step: 320, loss: 7.004759390838444e-05
step: 330, loss: 0.005075342021882534
step: 340, loss: 0.0005406843265518546
step: 350, loss: 0.00033930130302906036
step: 360, loss: 0.00026086982688866556
step: 370, loss: 9.945653437171131e-05
step: 380, loss: 0.0010539772920310497
step: 390, loss: 0.010888975113630295
step: 400, loss: 9.556891745887697e-05
step: 410, loss: 0.0061929416842758656
step: 420, loss: 0.011001824401319027
step: 430, loss: 0.00020068195590283722
step: 440, loss: 0.0025521935895085335
step: 450, loss: 0.0024825904984027147
step: 460, loss: 0.0003233302559237927
step: 470, loss: 0.000183278345502913
step: 480, loss: 0.0002579485299065709
step: 490, loss: 6.389322516042739e-05
step: 500, loss: 6.935115379747003e-05
step: 510, loss: 0.00025663842097856104
step: 520, loss: 0.0002396057971054688
step: 530, loss: 6.344404391711578e-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=0.912882298424467, f1=0.8956197576887232,
best_f1=0.9034965034965035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.8578902351437137e-05
step: 10, loss: 0.0004827669763471931
step: 20, loss: 0.0001481641666032374

```

step: 30, loss: 0.0003507582878228277
step: 40, loss: 0.11316309869289398
step: 50, loss: 6.1002512666163966e-05
step: 60, loss: 3.815297895926051e-05
step: 70, loss: 0.0012344607384875417
step: 80, loss: 5.966416938463226e-05
step: 90, loss: 0.000831498415209353
step: 100, loss: 0.0010408585658296943
step: 110, loss: 0.005102951545268297
step: 120, loss: 0.00011696251749526709
step: 130, loss: 0.006030930206179619
step: 140, loss: 0.00017752678832039237
step: 150, loss: 0.00013363728066906333
step: 160, loss: 0.00037693994818255305
step: 170, loss: 0.001272568479180336
step: 180, loss: 6.534404383273795e-05
step: 190, loss: 0.0002234018757008016
step: 200, loss: 0.0002801746013574302
step: 210, loss: 0.004022262990474701
step: 220, loss: 8.011861064005643e-05
step: 230, loss: 0.00038769186357967556
step: 240, loss: 9.240317740477622e-05
step: 250, loss: 0.00022509740665555
step: 260, loss: 0.02222595550119877
step: 270, loss: 0.012571834027767181
step: 280, loss: 2.8013133487547748e-05
step: 290, loss: 6.692624447168782e-05
step: 300, loss: 3.571989145711996e-05
step: 310, loss: 0.001971642253920436
step: 320, loss: 0.00022466132941190153
step: 330, loss: 0.00045274387230165303
step: 340, loss: 0.00040846518822945654
step: 350, loss: 8.288903336506337e-05
step: 360, loss: 0.001694018254056573
step: 370, loss: 5.19565655849874e-05
step: 380, loss: 0.0017517601372674108
step: 390, loss: 0.0026436843909323215
step: 400, loss: 0.00023063740809448063
step: 410, loss: 0.0008156555704772472
step: 420, loss: 0.000478476780699566
step: 430, loss: 0.00015569893002975732
step: 440, loss: 0.00024043403391260654
step: 450, loss: 4.691313733928837e-05
step: 460, loss: 0.0005203045438975096
step: 470, loss: 0.00010350149386795238
step: 480, loss: 9.750796016305685e-05
step: 490, loss: 9.374100773129612e-05
step: 500, loss: 0.000523418071679771


```

step: 510, loss: 0.0015809816541150212
step: 520, loss: 7.126281707314774e-05
step: 530, loss: 0.0004749197978526354
/usr/local/lib/python3.7/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=0.9144434222631096, f1=0.8986175115207373,
best_f1=0.9034965034965035
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:29, 192.09it/s]
load_f1 = 0.9134300421151146
real_f1 = 0.9121495327102803
0it [00:00, ?it/s]/usr/local/lib/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, 189.31it/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 bert-base-uncased \
  --fp16 \
  --save_model

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

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

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

Defaults for this optimization level are:

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

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

```

After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5537285208702087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3846153846153846, 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: 0.503345787525177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.43478260869565216, f1=0.5294117647058824,
best_f1=0.5294117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4230138063430786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.46428571428571436, f1=0.3939393939393939,
best_f1=0.3939393939393939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35748857259750366
/usr/local/lib/python3.7/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=0.3333333333333333, f1=0.4324324324324324,
best_f1=0.3939393939393939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3738439977169037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5098039215686275, f1=0.4905660377358491,
best_f1=0.4905660377358491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2785224914550781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5217391304347826, 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: 0.2556169927120209
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5714285714285714, 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: 0.08991910517215729
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.5, f1=0.4615384615384615, 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: 0.004510669037699699
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6250000000000001, f1=0.6153846153846153,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1470484435558319
/usr/local/lib/python3.7/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=0.5925925925925927, f1=0.6470588235294117,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0342162661254406
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.689655172413793,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025465503334999084
/usr/local/lib/python3.7/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=0.5454545454545454, f1=0.6666666666666666,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.034106917679309845
/usr/local/lib/python3.7/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=0.6060606060606061, f1=0.5945945945945946,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018168674781918526
/usr/local/lib/python3.7/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=0.6060606060606061, f1=0.5945945945945946,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040347546339035034
/usr/local/lib/python3.7/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=0.6060606060606061, f1=0.5945945945945946,
best_f1=0.6153846153846153
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16

```

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 134489.66it/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.")
load_f1 = 0.6250000000000001
real_f1 = 0.6250000000000001
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 272.68it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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


```

rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
  UserWarning)
step: 0, loss: 0.6389184594154358
step: 10, loss: 0.6617250442504883
step: 20, loss: 0.48337918519973755
step: 30, loss: 0.38339993357658386
step: 40, loss: 0.1523703783750534
step: 50, loss: 0.04787345230579376
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.06595179438591003
step: 70, loss: 0.07814402878284454
step: 80, loss: 0.1384209543466568
step: 90, loss: 0.05373376980423927
step: 100, loss: 0.008892862126231194
step: 110, loss: 0.09345519542694092
step: 120, loss: 0.021016893908381462
step: 130, loss: 0.049464982002973557
step: 140, loss: 0.007662048563361168
step: 150, loss: 0.047531675547361374
step: 160, loss: 0.010834548622369766
step: 170, loss: 0.028583776205778122
step: 180, loss: 0.16198502480983734
step: 190, loss: 0.02003425732254982
step: 200, loss: 0.00976658146828413
step: 210, loss: 0.004992729984223843
step: 220, loss: 0.01361821498721838
step: 230, loss: 0.034417033195495605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9556313993174061, f1=0.9519450800915331,
best_f1=0.9519450800915331
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04342492297291756
step: 10, loss: 0.006201000884175301
step: 20, loss: 0.0714723989367485
step: 30, loss: 0.1712387651205063
step: 40, loss: 0.04957818239927292
step: 50, loss: 0.004125166218727827
step: 60, loss: 0.004797420464456081
step: 70, loss: 0.20010517537593842

```

```

step: 80, loss: 0.1142694354057312
step: 90, loss: 0.009328545071184635
step: 100, loss: 0.010786606930196285
step: 110, loss: 0.05352421849966049
step: 120, loss: 0.004157153889536858
step: 130, loss: 0.0360470674932003
step: 140, loss: 0.013308316469192505
step: 150, loss: 0.01267241407185793
step: 160, loss: 0.10497376322746277
step: 170, loss: 0.008626987226307392
step: 180, loss: 0.004909321665763855
step: 190, loss: 0.02608383633196354
step: 200, loss: 0.009319483302533627
step: 210, loss: 0.0008947817841544747
step: 220, loss: 0.021810313686728477
step: 230, loss: 0.08747311681509018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830890642615557, f1=0.9749430523917996,
best_f1=0.9749430523917996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009413440711796284
step: 10, loss: 0.008040261454880238
step: 20, loss: 0.0033978333231061697
step: 30, loss: 0.0018353326013311744
step: 40, loss: 0.11347100138664246
step: 50, loss: 0.021666062995791435
step: 60, loss: 0.00113878701813519
step: 70, loss: 0.015448889695107937
step: 80, loss: 0.0011711426777765155
step: 90, loss: 0.1493402123451233
step: 100, loss: 0.00219991453923285
step: 110, loss: 0.004374295007437468
step: 120, loss: 0.015516545623540878
step: 130, loss: 0.01585960015654564
step: 140, loss: 0.04801279306411743
step: 150, loss: 0.002550339326262474
step: 160, loss: 0.004030116833746433
step: 170, loss: 0.003906077705323696
step: 180, loss: 0.011536824516952038
step: 190, loss: 0.0181549321860075

```

```

step: 200, loss: 0.008764710277318954
step: 210, loss: 0.0009177479660138488
step: 220, loss: 0.0008660274324938655
step: 230, loss: 0.0005470538162626326
/usr/local/lib/python3.7/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=0.9767441860465117, f1=0.9732142857142857,
best_f1=0.9749430523917996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035508659202605486
step: 10, loss: 0.00029923825059086084
step: 20, loss: 0.0004901661886833608
step: 30, loss: 0.0007450731936842203
step: 40, loss: 0.011375986970961094
step: 50, loss: 0.001070691621862352
step: 60, loss: 0.02579382248222828
step: 70, loss: 0.000510183977894485
step: 80, loss: 0.02073846198618412
step: 90, loss: 0.0012855305103585124
step: 100, loss: 0.06270480155944824
step: 110, loss: 0.0018800720572471619
step: 120, loss: 0.0195096954703331
step: 130, loss: 0.0025301810819655657
step: 140, loss: 0.016325898468494415
step: 150, loss: 0.015139742754399776
step: 160, loss: 0.06509371101856232
step: 170, loss: 0.0458909310400486
step: 180, loss: 0.0013497619656845927
step: 190, loss: 0.010489738546311855
step: 200, loss: 0.0011253474513068795
step: 210, loss: 0.036591943353414536
step: 220, loss: 0.0012819498078897595
step: 230, loss: 0.003678193548694253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9840546697038726, f1=0.9749430523917996,
best_f1=0.9749430523917996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0005653771222569048
step: 10, loss: 0.0019333361415192485
step: 20, loss: 0.0034225862473249435
step: 30, loss: 0.0020621842704713345
step: 40, loss: 0.014157694764435291
step: 50, loss: 0.00033659671316854656
step: 60, loss: 0.002276303479447961
step: 70, loss: 0.02228379249572754
step: 80, loss: 0.0009973846608772874
step: 90, loss: 0.006276021711528301
step: 100, loss: 0.0011512661585584283
step: 110, loss: 0.0012259874492883682
step: 120, loss: 0.0004001940251328051
step: 130, loss: 0.0037496532313525677
step: 140, loss: 0.006379581987857819
step: 150, loss: 0.030887186527252197
step: 160, loss: 0.0008523687138222158
step: 170, loss: 0.008442881517112255
step: 180, loss: 0.07050206512212753
step: 190, loss: 0.0007886509411036968
step: 200, loss: 0.09650193899869919
step: 210, loss: 0.0012647738913074136
step: 220, loss: 0.011643406003713608
step: 230, loss: 0.0006184933008626103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865168539325843, f1=0.9775784753363228,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001247528474777937
step: 10, loss: 0.00047569014714099467
step: 20, loss: 0.0018200945341959596
step: 30, loss: 0.0008307772222906351
step: 40, loss: 0.0006136991432867944
step: 50, loss: 0.000210209094802849
step: 60, loss: 0.00034104479709640145
step: 70, loss: 0.00022240735415834934
step: 80, loss: 0.00046171084977686405
step: 90, loss: 0.010422125458717346
step: 100, loss: 0.07643965631723404
step: 110, loss: 0.0008965259185060859

```

```

step: 120, loss: 0.00027439265977591276
step: 130, loss: 0.004785190336406231
step: 140, loss: 0.00023197052360046655
step: 150, loss: 0.006487716920673847
step: 160, loss: 0.03775160387158394
step: 170, loss: 0.00024496507830917835
step: 180, loss: 0.017958611249923706
step: 190, loss: 0.010134967975318432
step: 200, loss: 0.004735399503260851
step: 210, loss: 0.005091678351163864
step: 220, loss: 0.000358696561306715
step: 230, loss: 0.1100442036986351
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9774266365688488,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0041060615330934525
step: 10, loss: 0.0001441210333723575
step: 20, loss: 0.00026321274344809353
step: 30, loss: 0.001723736640997231
step: 40, loss: 0.00012842340220231563
step: 50, loss: 0.0002156613045372069
step: 60, loss: 0.002301787491887808
step: 70, loss: 0.0029175400268286467
step: 80, loss: 0.00016099598724395037
step: 90, loss: 5.848207001690753e-05
step: 100, loss: 8.794127643341199e-05
step: 110, loss: 0.013886183500289917
step: 120, loss: 7.767023635096848e-05
step: 130, loss: 0.0001232721988344565
step: 140, loss: 0.000276696402579546
step: 150, loss: 0.0010105398250743747
step: 160, loss: 0.06066618114709854
step: 170, loss: 0.00026264548068866134
step: 180, loss: 0.00029092031763866544
step: 190, loss: 0.0004611910553649068
step: 200, loss: 0.006417838390916586
step: 210, loss: 7.612485933350399e-05
step: 220, loss: 0.0002464957651682198
step: 230, loss: 0.002561919391155243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9898534385569334, f1=0.9807909604519773,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001765219058142975
step: 10, loss: 0.006996434647589922
step: 20, loss: 0.0006693355389870703
step: 30, loss: 0.00016267220780719072
step: 40, loss: 0.0003260948578827083
step: 50, loss: 0.0011698442976921797
step: 60, loss: 0.0004016877501271665
step: 70, loss: 0.0003262005338910967
step: 80, loss: 0.00018686139083001763
step: 90, loss: 0.00023475312627851963
step: 100, loss: 0.00012875105312559754
step: 110, loss: 0.00013554308679886162
step: 120, loss: 0.0074895229190588
step: 130, loss: 0.0019415834685787559
step: 140, loss: 0.00010853923595277593
step: 150, loss: 0.0002710046246647835
step: 160, loss: 0.00011982230353169143
step: 170, loss: 7.416187872877344e-05
step: 180, loss: 0.00024069714709185064
step: 190, loss: 0.00011372831795597449
step: 200, loss: 0.0004890659474767745
step: 210, loss: 0.0001513638999313116
step: 220, loss: 0.00022892901324667037
step: 230, loss: 0.0003222145896870643
/usr/local/lib/python3.7/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=0.9853107344632768, f1=0.9773242630385486,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.336303340503946e-05
step: 10, loss: 0.0002903903659898788
step: 20, loss: 0.0001976394996745512
step: 30, loss: 8.690721006132662e-05
step: 40, loss: 0.016471097245812416
step: 50, loss: 0.00027710196445696056
step: 60, loss: 0.0001552088069729507
step: 70, loss: 0.0004686036263592541
```

```

step: 80, loss: 0.000803487200755626
step: 90, loss: 7.98079781816341e-05
step: 100, loss: 0.07819847762584686
step: 110, loss: 5.715682345908135e-05
step: 120, loss: 0.0002366385597269982
step: 130, loss: 0.12462932616472244
step: 140, loss: 0.021235719323158264
step: 150, loss: 0.001995289698243141
step: 160, loss: 9.062284516403452e-05
step: 170, loss: 8.38524429127574e-05
step: 180, loss: 0.02364582009613514
step: 190, loss: 0.002138395793735981
step: 200, loss: 6.319962267298251e-05
step: 210, loss: 6.781492265872657e-05
step: 220, loss: 0.000674273120239377
step: 230, loss: 0.00014437397476285696
/usr/local/lib/python3.7/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=0.9842342342342343, f1=0.9707207207207207,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013115776237100363
step: 10, loss: 0.00015960237942636013
step: 20, loss: 0.0005078121903352439
step: 30, loss: 8.745226659812033e-05
step: 40, loss: 0.00018342857947573066
step: 50, loss: 0.00011191795056220144
step: 60, loss: 0.00021720088261645287
step: 70, loss: 0.0003753102500922978
step: 80, loss: 9.523233893560246e-05
step: 90, loss: 0.0007635020883753896
step: 100, loss: 0.00017337824101559818
step: 110, loss: 0.00013429162208922207
step: 120, loss: 0.00011178467684658244
step: 130, loss: 6.852267688373104e-05
step: 140, loss: 0.05397268012166023
step: 150, loss: 0.002786267315968871
step: 160, loss: 3.8331447285600007e-05
step: 170, loss: 8.145047468133271e-05
step: 180, loss: 0.00027522165328264236
step: 190, loss: 0.05045264586806297
step: 200, loss: 3.784741056733765e-05
step: 210, loss: 6.14127202425152e-05
step: 220, loss: 0.0008105204324238002
step: 230, loss: 0.00011308400280540809

```

```

/usr/local/lib/python3.7/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=0.9787234042553192, f1=0.9753363228699552,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07976101338863373
step: 10, loss: 5.5373697250615805e-05
step: 20, loss: 0.006650098133832216
step: 30, loss: 0.007419077213853598
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.004796585999429226
step: 50, loss: 8.560967398807406e-05
step: 60, loss: 0.002886236412450671
step: 70, loss: 0.00020500508253462613
step: 80, loss: 0.00011137709225295112
step: 90, loss: 0.00011498968524392694
step: 100, loss: 0.00018327395082451403
step: 110, loss: 5.590174259850755e-05
step: 120, loss: 6.869190838187933e-05
step: 130, loss: 6.0401605878723785e-05
step: 140, loss: 6.611904973397031e-05
step: 150, loss: 0.0034398760180920362
step: 160, loss: 6.713598850183189e-05
step: 170, loss: 0.0035651924554258585
step: 180, loss: 6.407814362319186e-05
step: 190, loss: 5.45689654245507e-05
step: 200, loss: 5.6635555665707216e-05
step: 210, loss: 8.870998135535046e-05
step: 220, loss: 5.666624929290265e-05
step: 230, loss: 6.743729318259284e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9818594104308391, f1=0.9750566893424036,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.9869296819670126e-05
step: 10, loss: 9.465343464398757e-05
step: 20, loss: 7.835251017240807e-05
step: 30, loss: 9.288513683713973e-05
step: 40, loss: 0.004413362592458725
step: 50, loss: 0.0062937261536717415
step: 60, loss: 0.008744166232645512

```



```

step: 70, loss: 0.00010875705629587173
step: 80, loss: 0.0004339797014836222
step: 90, loss: 8.284511568490416e-05
step: 100, loss: 7.369406375801191e-05
step: 110, loss: 5.825351399835199e-05
step: 120, loss: 8.266569784609601e-05
step: 130, loss: 6.862147711217403e-05
step: 140, loss: 5.556667747441679e-05
step: 150, loss: 0.00046874943654984236
step: 160, loss: 6.198125629452989e-05
step: 170, loss: 0.00010009611287387088
step: 180, loss: 8.170148066710681e-05
step: 190, loss: 0.005652001593261957
step: 200, loss: 3.5787590604741126e-05
step: 210, loss: 5.1674633141374215e-05
step: 220, loss: 5.763195440522395e-05
step: 230, loss: 7.717539847362787e-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=0.9842342342342343, f1=0.9750566893424036,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.339982352685183e-05
step: 10, loss: 0.0017141588032245636
step: 20, loss: 0.002276371465995908
step: 30, loss: 7.637315866304561e-05
step: 40, loss: 0.00045032487832941115
step: 50, loss: 0.00012830027844756842
step: 60, loss: 7.422352064168081e-05
step: 70, loss: 3.6761968658538535e-05
step: 80, loss: 5.262708509690128e-05
step: 90, loss: 0.00016018228780012578
step: 100, loss: 2.6858622732106596e-05
step: 110, loss: 0.008941391482949257
step: 120, loss: 0.0017220211448147893
step: 130, loss: 0.00013180356472730637
step: 140, loss: 3.9829101297073066e-05
step: 150, loss: 3.833919618045911e-05
step: 160, loss: 0.000795534928329289
step: 170, loss: 8.082784188445657e-05
step: 180, loss: 5.444525959319435e-05
step: 190, loss: 0.00014389125863090158
step: 200, loss: 0.006490643136203289
step: 210, loss: 3.25170076393988e-05
step: 220, loss: 5.68238647247199e-05

```

```

step: 230, loss: 4.344909393694252e-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=0.987598647125141, f1=0.9774266365688488,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.0257800719700754e-05
step: 10, loss: 0.006632282864302397
step: 20, loss: 0.00013310274516697973
step: 30, loss: 0.00026719667948782444
step: 40, loss: 0.0020047163125127554
step: 50, loss: 6.460247823270038e-05
step: 60, loss: 0.00016900515765883029
step: 70, loss: 4.5059070544084534e-05
step: 80, loss: 4.397139491629787e-05
step: 90, loss: 4.210532279103063e-05
step: 100, loss: 4.4179421820444986e-05
step: 110, loss: 0.0001408426760463044
step: 120, loss: 2.239228706457652e-05
step: 130, loss: 9.364934521727264e-05
step: 140, loss: 0.0003689866862259805
step: 150, loss: 3.934407141059637e-05
step: 160, loss: 0.03207571059465408
step: 170, loss: 5.651870014844462e-05
step: 180, loss: 3.9895749068818986e-05
step: 190, loss: 5.1912029448430985e-05
step: 200, loss: 6.569344986928627e-05
step: 210, loss: 0.0001815089926822111
step: 220, loss: 0.00018044895841740072
step: 230, loss: 0.0001359035522909835
/usr/local/lib/python3.7/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=0.9863636363636363, f1=0.9738339021615473,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.558643948053941e-05
step: 10, loss: 3.437103077885695e-05
step: 20, loss: 4.4525921111926436e-05
step: 30, loss: 7.057406764943153e-05
step: 40, loss: 0.0003096394066233188
step: 50, loss: 0.021272175014019012
step: 60, loss: 3.9575894334120676e-05

```

```

step: 70, loss: 6.447229679906741e-05
step: 80, loss: 0.0001183835047413595
step: 90, loss: 4.697397889685817e-05
step: 100, loss: 0.004011720884591341
step: 110, loss: 9.431866783415899e-05
step: 120, loss: 7.638841634616256e-05
step: 130, loss: 3.313176421215758e-05
step: 140, loss: 3.346307858009823e-05
step: 150, loss: 0.00011941747652599588
step: 160, loss: 0.0001648177276365459
step: 170, loss: 4.24528589064721e-05
step: 180, loss: 4.9393591325497255e-05
step: 190, loss: 5.341726136975922e-05
step: 200, loss: 4.3579402699833736e-05
step: 210, loss: 4.430290209711529e-05
step: 220, loss: 3.7679328670492396e-05
step: 230, loss: 5.075382432551123e-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=0.9863636363636363, f1=0.9726651480637813,
best_f1=0.9807909604519773
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 197.52it/s]
load_f1 = 0.9863945578231292
real_f1 = 0.9864253393665158
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 258.08it/s]

```

2.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

step: 0, loss: 0.6421183943748474

step: 10, loss: 0.6995790600776672

step: 20, loss: 0.6058841347694397

step: 30, loss: 0.3197649419307709

step: 40, loss: 0.2713649868965149

step: 50, loss: 0.2516390085220337

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

step: 60, loss: 0.679408848285675

step: 70, loss: 0.27800214290618896

step: 80, loss: 0.052117928862571716

step: 90, loss: 0.424496054649353

step: 100, loss: 0.11637584865093231

step: 110, loss: 0.1040760725736618

step: 120, loss: 0.16227833926677704

step: 130, loss: 0.1160588189959526

```

step: 140, loss: 0.16380897164344788
step: 150, loss: 0.18809163570404053
step: 160, loss: 0.04103270545601845
step: 170, loss: 0.26856333017349243
step: 180, loss: 0.1332913488149643
step: 190, loss: 0.05380196124315262
step: 200, loss: 0.14468629658222198
step: 210, loss: 0.06936771422624588
step: 220, loss: 0.26232823729515076
step: 230, loss: 0.20701025426387787
step: 240, loss: 0.11325754225254059
step: 250, loss: 0.09236881136894226
step: 260, loss: 0.41953715682029724
step: 270, loss: 0.043145932257175446
step: 280, loss: 0.09062140434980392
step: 290, loss: 0.0857018232345581
step: 300, loss: 0.0630778893828392
step: 310, loss: 0.2225126177072525
step: 320, loss: 0.1418837010860443
step: 330, loss: 0.06267087161540985
step: 340, loss: 0.1708649843931198
step: 350, loss: 0.054169122129678726
step: 360, loss: 0.036410946398973465
step: 370, loss: 0.16299636662006378
step: 380, loss: 0.012924149632453918
step: 390, loss: 0.3910031318664551
step: 400, loss: 0.2679772675037384
step: 410, loss: 0.06670437008142471
step: 420, loss: 0.10219503939151764
step: 430, loss: 0.20843036472797394
step: 440, loss: 0.059280648827552795
step: 450, loss: 0.041594695299863815
step: 460, loss: 0.024200303480029106
step: 470, loss: 0.04488743841648102
step: 480, loss: 0.14717774093151093
step: 490, loss: 0.3019404113292694
step: 500, loss: 0.10322026908397675
step: 510, loss: 0.14267446100711823
step: 520, loss: 0.031807780265808105
step: 530, loss: 0.01412262860685587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8805206880520688, f1=0.8839981659789088,

```

```
best_f1=0.8839981659789088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20008733868598938
step: 10, loss: 0.14765706658363342
step: 20, loss: 0.01554082054644823
step: 30, loss: 0.014959963969886303
step: 40, loss: 0.07016542553901672
step: 50, loss: 0.17874997854232788
step: 60, loss: 0.029129650443792343
step: 70, loss: 0.06266877055168152
step: 80, loss: 0.1823686808347702
step: 90, loss: 0.09609533101320267
step: 100, loss: 0.1751047670841217
step: 110, loss: 0.058328062295913696
step: 120, loss: 0.13430652022361755
step: 130, loss: 0.09947343170642853
step: 140, loss: 0.0451570563018322
step: 150, loss: 0.08646640181541443
step: 160, loss: 0.10795057564973831
step: 170, loss: 0.06261748820543289
step: 180, loss: 0.03127169981598854
step: 190, loss: 0.027891673147678375
step: 200, loss: 0.04309195280075073
step: 210, loss: 0.03841289505362511
step: 220, loss: 0.09594997018575668
step: 230, loss: 0.011061075143516064
step: 240, loss: 0.03179028630256653
step: 250, loss: 0.02923828735947609
step: 260, loss: 0.004542454611510038
step: 270, loss: 0.24263595044612885
step: 280, loss: 0.10136362165212631
step: 290, loss: 0.13415347039699554
step: 300, loss: 0.03077983483672142
step: 310, loss: 0.08932148665189743
step: 320, loss: 0.24387158453464508
step: 330, loss: 0.11128108203411102
step: 340, loss: 0.01271548867225647
step: 350, loss: 0.0038324170745909214
step: 360, loss: 0.14977505803108215
step: 370, loss: 0.1294868290424347
step: 380, loss: 0.18478482961654663
step: 390, loss: 0.1024661585688591
step: 400, loss: 0.061028625816106796
step: 410, loss: 0.015615259297192097
step: 420, loss: 0.04466495290398598
step: 430, loss: 0.10303568094968796
```

```

step: 440, loss: 0.0738116130232811
step: 450, loss: 0.07059214264154434
step: 460, loss: 0.026717405766248703
step: 470, loss: 0.06706325709819794
step: 480, loss: 0.16970719397068024
step: 490, loss: 0.04030390456318855
step: 500, loss: 0.31226709485054016
step: 510, loss: 0.030290530994534492
step: 520, loss: 0.03405417129397392
step: 530, loss: 0.022001171484589577
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9136385087305333, f1=0.9110386585933862,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11133322864770889
step: 10, loss: 0.0314088799059391
step: 20, loss: 0.11916080117225647
step: 30, loss: 0.08529753237962723
step: 40, loss: 0.012435915879905224
step: 50, loss: 0.17156511545181274
step: 60, loss: 0.014752830378711224
step: 70, loss: 0.02679400146007538
step: 80, loss: 0.02188771590590477
step: 90, loss: 0.0072515239007771015
step: 100, loss: 0.025901200249791145
step: 110, loss: 0.0186820887029171
step: 120, loss: 0.03384280204772949
step: 130, loss: 0.006153373047709465
step: 140, loss: 0.017642298713326454
step: 150, loss: 0.03231294825673103
step: 160, loss: 0.04689835384488106
step: 170, loss: 0.1474829614162445
step: 180, loss: 0.02719896286725998
step: 190, loss: 0.12522439658641815
step: 200, loss: 0.07341361045837402
step: 210, loss: 0.02814071998000145
step: 220, loss: 0.19755366444587708
step: 230, loss: 0.0656396672129631
step: 240, loss: 0.10897857695817947
step: 250, loss: 0.11751477420330048

```



```

step: 260, loss: 0.02231142483651638
step: 270, loss: 0.014808163978159428
step: 280, loss: 0.1306544989347458
step: 290, loss: 0.004855015780776739
step: 300, loss: 0.06780495494604111
step: 310, loss: 0.06948604434728622
step: 320, loss: 0.033381327986717224
step: 330, loss: 0.0187408234924078
step: 340, loss: 0.018932919949293137
step: 350, loss: 0.005653982516378164
step: 360, loss: 0.06365828961133957
step: 370, loss: 0.017108822241425514
step: 380, loss: 0.051862411201000214
step: 390, loss: 0.1518184095621109
step: 400, loss: 0.08538372069597244
step: 410, loss: 0.012913231737911701
step: 420, loss: 0.12404289841651917
step: 430, loss: 0.06413472443819046
step: 440, loss: 0.021685147657990456
step: 450, loss: 0.06255479902029037
step: 460, loss: 0.02820744179189205
step: 470, loss: 0.10914332419633865
step: 480, loss: 0.05539077892899513
step: 490, loss: 0.016409706324338913
step: 500, loss: 0.055109284818172455
step: 510, loss: 0.016850585117936134
step: 520, loss: 0.23423655331134796
step: 530, loss: 0.0472731739282608
/usr/local/lib/python3.7/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=0.9064546304957904, f1=0.904363974001857,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0176222026348114
step: 10, loss: 0.014656643383204937
step: 20, loss: 0.002382960868999362
step: 30, loss: 0.004833536688238382
step: 40, loss: 0.010605007410049438
step: 50, loss: 0.0459066741168499
step: 60, loss: 0.030005773529410362
step: 70, loss: 0.03944934904575348
step: 80, loss: 0.04000858590006828
step: 90, loss: 0.17409229278564453
step: 100, loss: 0.09159214049577713
step: 110, loss: 0.1408403068780899

```

```
step: 120, loss: 0.040494050830602646
step: 130, loss: 0.07796952873468399
step: 140, loss: 0.0033645813819020987
step: 150, loss: 0.007714849431067705
step: 160, loss: 0.1998731940984726
step: 170, loss: 0.054740842431783676
step: 180, loss: 0.014317610301077366
step: 190, loss: 0.027514822781085968
step: 200, loss: 0.158567413687706
step: 210, loss: 0.11601319164037704
step: 220, loss: 0.020718451589345932
step: 230, loss: 0.030956992879509926
step: 240, loss: 0.013415921479463577
step: 250, loss: 0.08537545800209045
step: 260, loss: 0.010814519599080086
step: 270, loss: 0.03729984909296036
step: 280, loss: 0.13037899136543274
step: 290, loss: 0.019299915060400963
step: 300, loss: 0.007722924463450909
step: 310, loss: 0.014600833877921104
step: 320, loss: 0.00975812878459692
step: 330, loss: 0.18026678264141083
step: 340, loss: 0.05549932271242142
step: 350, loss: 0.02072390355169773
step: 360, loss: 0.0885450541973114
step: 370, loss: 0.003625981044024229
step: 380, loss: 0.008933785371482372
step: 390, loss: 0.004333481192588806
step: 400, loss: 0.013652737252414227
step: 410, loss: 0.03366442024707794
step: 420, loss: 0.019414102658629417
step: 430, loss: 0.17957621812820435
step: 440, loss: 0.01042608916759491
step: 450, loss: 0.033339135348796844
step: 460, loss: 0.06916292011737823
step: 470, loss: 0.019884275272488594
step: 480, loss: 0.03323392942547798
step: 490, loss: 0.0013498755870386958
step: 500, loss: 0.012209564447402954
step: 510, loss: 0.02152399905025959
step: 520, loss: 0.06027406454086304
step: 530, loss: 0.015798967331647873
/usr/local/lib/python3.7/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=0.9005090236001851, f1=0.8991674375578168,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 0.041906509548425674
step: 10, loss: 0.07248084992170334
step: 20, loss: 0.006297219078987837
step: 30, loss: 0.00432035094127059
step: 40, loss: 0.06559421867132187
step: 50, loss: 0.005774305202066898
step: 60, loss: 0.07638311386108398
step: 70, loss: 0.008964503183960915
step: 80, loss: 0.0009346548467874527
step: 90, loss: 0.07939144223928452
step: 100, loss: 0.0028958767652511597
step: 110, loss: 0.009304646402597427
step: 120, loss: 0.007788783870637417
step: 130, loss: 0.0010242164134979248
step: 140, loss: 0.008586565032601357
step: 150, loss: 0.0011852829484269023
step: 160, loss: 0.007403763942420483
step: 170, loss: 0.006036612670868635
step: 180, loss: 0.0021355533972382545
step: 190, loss: 0.0020545825827866793
step: 200, loss: 0.0011980661656707525
step: 210, loss: 0.00880391988903284
step: 220, loss: 0.04685370996594429
step: 230, loss: 0.00364682893268764
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 240, loss: 0.0015446157194674015
step: 250, loss: 0.007418077904731035
step: 260, loss: 0.026709582656621933
step: 270, loss: 0.03802258148789406
step: 280, loss: 0.010731329210102558
step: 290, loss: 0.125783309340477
step: 300, loss: 0.020489608868956566
step: 310, loss: 0.004768455866724253
step: 320, loss: 0.028641507029533386
step: 330, loss: 0.005671055056154728
step: 340, loss: 0.0023975130170583725
step: 350, loss: 0.0007616751827299595
step: 360, loss: 0.032008878886699677
step: 370, loss: 0.0559818372130394
step: 380, loss: 0.0042708637192845345
step: 390, loss: 0.007283355109393597
step: 400, loss: 0.018623489886522293
step: 410, loss: 0.06791634857654572
step: 420, loss: 0.00822922307997942
step: 430, loss: 0.007047269959002733
step: 440, loss: 0.1229095309972763
```

```

step: 450, loss: 0.15813925862312317
step: 460, loss: 0.015805184841156006
step: 470, loss: 0.002564621390774846
step: 480, loss: 0.007043006829917431
step: 490, loss: 0.001240032841451466
step: 500, loss: 0.005430764053016901
step: 510, loss: 0.02720893733203411
step: 520, loss: 0.07667100429534912
step: 530, loss: 0.01044459082186222
/usr/local/lib/python3.7/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=0.893461362597165, f1=0.9012915129151292,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010381587781012058
step: 10, loss: 0.001397965126670897
step: 20, loss: 0.15348173677921295
step: 30, loss: 0.024692337960004807
step: 40, loss: 0.013222410343587399
step: 50, loss: 0.0006974615971557796
step: 60, loss: 0.0003243243554607034
step: 70, loss: 0.00028659726376645267
step: 80, loss: 0.0010723741725087166
step: 90, loss: 0.00016416743164882064
step: 100, loss: 0.003354474436491728
step: 110, loss: 0.03272305428981781
step: 120, loss: 0.0005049642059020698
step: 130, loss: 0.014998250640928745
step: 140, loss: 0.004873654805123806
step: 150, loss: 0.014294939115643501
step: 160, loss: 0.003355575492605567
step: 170, loss: 0.0005617736023850739
step: 180, loss: 0.0006329663447104394
step: 190, loss: 0.007124622818082571
step: 200, loss: 0.0026829682756215334
step: 210, loss: 0.009584969840943813
step: 220, loss: 0.02568047121167183
step: 230, loss: 0.0013828420778736472
step: 240, loss: 0.06795831769704819
step: 250, loss: 0.007625112775713205
step: 260, loss: 0.008885785937309265
step: 270, loss: 0.05638045072555542
step: 280, loss: 0.0030669791158288717
step: 290, loss: 0.0023569113109260798
step: 300, loss: 0.005271392408758402

```

```

step: 310, loss: 0.004115608055144548
step: 320, loss: 0.00020564069563988596
step: 330, loss: 0.0003112927370239049
step: 340, loss: 0.0010278798872604966
step: 350, loss: 0.0015038878191262484
step: 360, loss: 0.007017635740339756
step: 370, loss: 0.015463240444660187
step: 380, loss: 0.0019163333345204592
step: 390, loss: 0.024501746520400047
step: 400, loss: 0.04354168102145195
step: 410, loss: 0.002379309618845582
step: 420, loss: 0.0028270878829061985
step: 430, loss: 0.0010832464322447777
step: 440, loss: 0.014979901723563671
step: 450, loss: 0.003525958862155676
step: 460, loss: 0.005017329938709736
step: 470, loss: 0.037662573158741
step: 480, loss: 0.003897111862897873
step: 490, loss: 0.012422401458024979
step: 500, loss: 0.005931271705776453
step: 510, loss: 0.003577057272195816
step: 520, loss: 0.0107789421454072
step: 530, loss: 0.10782032459974289
/usr/local/lib/python3.7/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=0.9013953488372093, f1=0.8959107806691451,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012953934259712696
step: 10, loss: 0.003229910973459482
step: 20, loss: 0.001477081561461091
step: 30, loss: 0.0007440935587510467
step: 40, loss: 0.0013638341333717108
step: 50, loss: 0.007639977615326643
step: 60, loss: 0.0017430043080821633
step: 70, loss: 0.0008977458928711712
step: 80, loss: 0.0016804892802610993
step: 90, loss: 0.009531316347420216
step: 100, loss: 0.0571146197617054
step: 110, loss: 0.0004508172278292477
step: 120, loss: 0.006238567642867565
step: 130, loss: 0.005288966931402683
step: 140, loss: 0.000607315159868449
step: 150, loss: 0.0010030283592641354
step: 160, loss: 0.0002614394179545343

```

```

step: 170, loss: 0.031872231513261795
step: 180, loss: 0.0003760058607440442
step: 190, loss: 0.0003110631078016013
step: 200, loss: 0.002712561748921871
step: 210, loss: 0.007078133523464203
step: 220, loss: 0.0021179141476750374
step: 230, loss: 0.004019244108349085
step: 240, loss: 0.03200826048851013
step: 250, loss: 0.0018088252982124686
step: 260, loss: 0.013130580075085163
step: 270, loss: 0.0013623659033328295
step: 280, loss: 0.003914074506610632
step: 290, loss: 0.015767047181725502
step: 300, loss: 0.0005235964199528098
step: 310, loss: 0.0007392773986794055
step: 320, loss: 0.0029235773254185915
step: 330, loss: 0.0007404144271276891
step: 340, loss: 0.0047920155338943005
step: 350, loss: 0.03492523729801178
step: 360, loss: 0.002458566101267934
step: 370, loss: 0.0013128904392942786
step: 380, loss: 0.00302308052778244
step: 390, loss: 0.00016593914187978953
step: 400, loss: 0.0012874723179265857
step: 410, loss: 0.0031609944999217987
step: 420, loss: 0.0001320415030932054
step: 430, loss: 7.420865586027503e-05
step: 440, loss: 0.06822704523801804
step: 450, loss: 0.0011051562614738941
step: 460, loss: 0.009024790488183498
step: 470, loss: 0.0008923602872528136
step: 480, loss: 0.1847154051065445
step: 490, loss: 0.07556650787591934
step: 500, loss: 0.00400968873873353
step: 510, loss: 0.011896847747266293
step: 520, loss: 0.00691658491268754
step: 530, loss: 0.024502204731106758
/usr/local/lib/python3.7/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=0.9050543221539914, f1=0.8994808872109485,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004701008554548025
step: 10, loss: 0.000747853540815413
step: 20, loss: 0.0002490698825567961

```

step: 30, loss: 0.005752027966082096
step: 40, loss: 0.0009827420581132174
step: 50, loss: 0.05269772931933403
step: 60, loss: 0.002372696530073881
step: 70, loss: 0.002257322194054723
step: 80, loss: 0.004629182629287243
step: 90, loss: 0.0005645881174132228
step: 100, loss: 0.0002573887468315661
step: 110, loss: 0.0001688704069238156
step: 120, loss: 0.0018263008678331971
step: 130, loss: 0.00015738687943667173
step: 140, loss: 0.026939496397972107
step: 150, loss: 0.0009381232084706426
step: 160, loss: 0.007757557090371847
step: 170, loss: 0.00883416272699833
step: 180, loss: 0.002959827659651637
step: 190, loss: 0.005395802669227123
step: 200, loss: 0.001978703774511814
step: 210, loss: 0.0013824740890413523
step: 220, loss: 0.001065244898200035
step: 230, loss: 0.0011386770056560636
step: 240, loss: 0.0016238264506682754
step: 250, loss: 9.096676512854174e-05
step: 260, loss: 5.8929534134222195e-05
step: 270, loss: 0.0002794179890770465
step: 280, loss: 0.0002789598365779966
step: 290, loss: 0.0002508426259737462
step: 300, loss: 0.000900502607692033
step: 310, loss: 0.005109009332954884
step: 320, loss: 0.06144782900810242
step: 330, loss: 8.4844978118781e-05
step: 340, loss: 5.4401421948568895e-05
step: 350, loss: 7.197360537247732e-05
step: 360, loss: 0.004556639585644007
step: 370, loss: 0.0004421162884682417
step: 380, loss: 0.0010050575947389007
step: 390, loss: 0.09228521585464478
step: 400, loss: 0.00038622028660029173
step: 410, loss: 7.893027941463515e-05
step: 420, loss: 0.0011400915682315826
step: 430, loss: 0.05881895497441292
step: 440, loss: 0.045497827231884
step: 450, loss: 0.029567716643214226
step: 460, loss: 0.007798568811267614
step: 470, loss: 0.0013905331725254655
step: 480, loss: 0.00018915344844572246
step: 490, loss: 0.0008220889721997082
step: 500, loss: 0.012848912738263607

```

step: 510, loss: 0.0005905778962187469
step: 520, loss: 0.00949020218104124
step: 530, loss: 0.0005724729271605611
/usr/local/lib/python3.7/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=0.9019607843137254, f1=0.8994857410004675,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015521423192694783
step: 10, loss: 0.0004743530007544905
step: 20, loss: 0.00039268372347578406
step: 30, loss: 0.0006150699337013066
step: 40, loss: 0.00018168757378589362
step: 50, loss: 0.0006120430771261454
step: 60, loss: 0.0001891387510113418
step: 70, loss: 0.0001761371677275747
step: 80, loss: 0.04920051246881485
step: 90, loss: 0.006606161128729582
step: 100, loss: 0.00011025486310245469
step: 110, loss: 7.953720341902226e-05
step: 120, loss: 0.00018086179625242949
step: 130, loss: 0.016220830380916595
step: 140, loss: 0.00016000738833099604
step: 150, loss: 8.190477092284709e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.0023670480586588383
step: 170, loss: 0.0120537793263793
step: 180, loss: 0.0005745809758082032
step: 190, loss: 7.674529479118064e-05
step: 200, loss: 0.00013205198047216982
step: 210, loss: 0.000986958504654467
step: 220, loss: 0.0025872173719108105
step: 230, loss: 0.00014576951798517257
step: 240, loss: 0.009466761723160744
step: 250, loss: 6.879949796712026e-05
step: 260, loss: 0.0034492406994104385
step: 270, loss: 8.215332491090521e-05
step: 280, loss: 8.99720544111915e-05
step: 290, loss: 0.1954873502254486
step: 300, loss: 0.0005471086478792131
step: 310, loss: 0.0425892248749733
step: 320, loss: 7.263939187396318e-05
step: 330, loss: 0.000770552025642246
step: 340, loss: 0.0004944056272506714
step: 350, loss: 0.014357835054397583

```



```

step: 360, loss: 6.871293589938432e-05
step: 370, loss: 6.76411873428151e-05
step: 380, loss: 0.12197710573673248
step: 390, loss: 9.527731162961572e-05
step: 400, loss: 0.022396525368094444
step: 410, loss: 8.096696546999738e-05
step: 420, loss: 0.0001123626425396651
step: 430, loss: 0.0001423616340616718
step: 440, loss: 0.001385149429552257
step: 450, loss: 6.076516001485288e-05
step: 460, loss: 0.00032199808629229665
step: 470, loss: 2.8888822271255776e-05
step: 480, loss: 8.181224256986752e-05
step: 490, loss: 0.015059621073305607
step: 500, loss: 0.0006363976281136274
step: 510, loss: 0.0002029921452049166
step: 520, loss: 0.0024405813310295343
step: 530, loss: 8.022625115700066e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9047619047619049, f1=0.9072635906806761,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005524276290088892
step: 10, loss: 4.131819150643423e-05
step: 20, loss: 3.266232306486927e-05
step: 30, loss: 3.914746775990352e-05
step: 40, loss: 0.006353192497044802
step: 50, loss: 4.2652460251701996e-05
step: 60, loss: 4.172064654994756e-05
step: 70, loss: 4.869724580203183e-05
step: 80, loss: 4.449639163794927e-05
step: 90, loss: 6.227455742191523e-05
step: 100, loss: 3.163423025398515e-05
step: 110, loss: 5.750155833084136e-05
step: 120, loss: 0.0029187360778450966
step: 130, loss: 0.0002801061200443655
step: 140, loss: 0.008070878684520721
step: 150, loss: 0.003944296855479479
step: 160, loss: 0.006311206612735987
step: 170, loss: 0.0013673445209860802
step: 180, loss: 0.0006554963765665889
step: 190, loss: 0.0065975673496723175
step: 200, loss: 0.01139528676867485
step: 210, loss: 0.0007435818552039564

```

```

step: 220, loss: 0.008809467777609825
step: 230, loss: 0.002367014065384865
step: 240, loss: 0.00029429158894345164
step: 250, loss: 0.0003326826263219118
step: 260, loss: 0.0017934024799615145
step: 270, loss: 0.00047761667519807816
step: 280, loss: 0.008278603665530682
step: 290, loss: 0.0002993041416630149
step: 300, loss: 3.826117972494103e-05
step: 310, loss: 5.3111394663574174e-05
step: 320, loss: 0.0018611373379826546
step: 330, loss: 9.12389878067188e-05
step: 340, loss: 0.0016140480292961001
step: 350, loss: 0.010330650955438614
step: 360, loss: 0.026632066816091537
step: 370, loss: 0.0021383147686719894
step: 380, loss: 0.014983500353991985
step: 390, loss: 0.0013607689179480076
step: 400, loss: 0.0014691282995045185
step: 410, loss: 0.00016818681615404785
step: 420, loss: 0.0019325728062540293
step: 430, loss: 6.737450894434005e-05
step: 440, loss: 0.012971889227628708
step: 450, loss: 0.00010554893378866836
step: 460, loss: 0.00039779514190740883
step: 470, loss: 0.0012703222455456853
step: 480, loss: 0.0006772240740247071
step: 490, loss: 0.0004273222293704748
step: 500, loss: 0.032655611634254456
step: 510, loss: 0.003163197310641408
step: 520, loss: 0.00010220407421002164
step: 530, loss: 8.157424599630758e-05
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 10: dev_f1=0.9049645390070921, f1=0.899624765478424,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008676313445903361
step: 10, loss: 0.0006851975340396166
step: 20, loss: 0.00016965520626399666
step: 30, loss: 0.002202762523666024
step: 40, loss: 0.01457873173058033
step: 50, loss: 0.00827951543033123
step: 60, loss: 0.10949289798736572

```

step: 70, loss: 0.00015370994515251368
step: 80, loss: 5.083488940726966e-05
step: 90, loss: 0.00014331706915982068
step: 100, loss: 0.00011411857849452645
step: 110, loss: 0.0007467767572961748
step: 120, loss: 0.008030190132558346
step: 130, loss: 0.0012494346592575312
step: 140, loss: 0.0002891597105190158
step: 150, loss: 0.0014395692851394415
step: 160, loss: 6.766614387743175e-05
step: 170, loss: 6.105282955104485e-05
step: 180, loss: 0.0008487323648296297
step: 190, loss: 0.0008385212859138846
step: 200, loss: 0.0002944681327790022
step: 210, loss: 0.00011318797623971477
step: 220, loss: 0.0009123371564783156
step: 230, loss: 0.00014548914623446763
step: 240, loss: 0.0001000093761831522
step: 250, loss: 6.51989466859959e-05
step: 260, loss: 3.435330290812999e-05
step: 270, loss: 0.00011668509978335351
step: 280, loss: 0.014529834501445293
step: 290, loss: 8.19670021883212e-05
step: 300, loss: 0.000226349409786053
step: 310, loss: 4.4794101995648816e-05
step: 320, loss: 5.499663529917598e-05
step: 330, loss: 9.474795660935342e-05
step: 340, loss: 6.289568409556523e-05
step: 350, loss: 0.00011626882769633085
step: 360, loss: 2.6128600438823923e-05
step: 370, loss: 0.0001762302708812058
step: 380, loss: 2.904539360315539e-05
step: 390, loss: 3.391002974240109e-05
step: 400, loss: 5.3535284678218886e-05
step: 410, loss: 0.00219617853872478
step: 420, loss: 0.00022536408505402505
step: 430, loss: 3.63798244507052e-05
step: 440, loss: 0.000312906748149544
step: 450, loss: 5.1432110922178254e-05
step: 460, loss: 0.07604539394378662
step: 470, loss: 3.69305198546499e-05
step: 480, loss: 0.0001205264707095921
step: 490, loss: 0.0007117890636436641
step: 500, loss: 0.00019556177721824497
step: 510, loss: 4.327644273871556e-05
step: 520, loss: 9.75517978076823e-05
step: 530, loss: 0.00022600029478780925
/usr/local/lib/python3.7/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=0.8963503649635037, f1=0.8793436293436293,  
best_f1=0.9110386585933862  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 7.525386172346771e-05  
step: 10, loss: 2.9443719540722668e-05  
step: 20, loss: 8.91874369699508e-05  
step: 30, loss: 3.5101598768960685e-05  
step: 40, loss: 0.006304269190877676  
step: 50, loss: 0.004015741404145956  
step: 60, loss: 0.00023556224186904728  
step: 70, loss: 0.0006701695965602994  
step: 80, loss: 2.7592037440626882e-05  
step: 90, loss: 8.17969412310049e-05  
step: 100, loss: 0.004596303217113018  
step: 110, loss: 5.970377605990507e-05  
step: 120, loss: 3.17305421049241e-05  
step: 130, loss: 0.001228134031407535  
step: 140, loss: 0.0014703803462907672  
step: 150, loss: 5.492682248586789e-05  
step: 160, loss: 6.0098936955910176e-05  
step: 170, loss: 8.654796693008393e-05  
step: 180, loss: 4.392215487314388e-05  
step: 190, loss: 4.141864337725565e-05  
step: 200, loss: 0.0025979606434702873  
step: 210, loss: 0.00012295546184759587  
step: 220, loss: 7.767311035422608e-05  
step: 230, loss: 3.231120717828162e-05  
step: 240, loss: 9.510204108664766e-05  
step: 250, loss: 3.4226744901388884e-05  
step: 260, loss: 4.916509715258144e-05  
step: 270, loss: 7.54372522351332e-05  
step: 280, loss: 0.0017708134837448597  
step: 290, loss: 0.0007976074703037739  
step: 300, loss: 0.0012471232330426574  
step: 310, loss: 0.000434413377661258  
step: 320, loss: 3.798383113462478e-05  
step: 330, loss: 6.382331775967032e-05  
step: 340, loss: 6.006129115121439e-05  
step: 350, loss: 0.0018881751457229257  
step: 360, loss: 0.00013040051271673292  
step: 370, loss: 0.0003307789156679064  
step: 380, loss: 8.8680608314462e-05  
step: 390, loss: 0.00025736266979947686  
step: 400, loss: 4.4820248149335384e-05
```

```

step: 410, loss: 0.00017747105448506773
step: 420, loss: 0.02805846929550171
step: 430, loss: 0.011089909821748734
step: 440, loss: 9.772148769116029e-05
step: 450, loss: 0.00022864325728733093
step: 460, loss: 4.317805723985657e-05
step: 470, loss: 0.00037116423482075334
step: 480, loss: 0.00042619919986464083
step: 490, loss: 0.014076890423893929
step: 500, loss: 3.85129897040315e-05
step: 510, loss: 0.0008898564847186208
step: 520, loss: 0.0002672789851203561
step: 530, loss: 6.75161209073849e-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=0.9096098953377736, f1=0.9006181645268663,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.530167204095051e-05
step: 10, loss: 0.0001024241209961474
step: 20, loss: 0.00042126711923629045
step: 30, loss: 5.974684972898103e-05
step: 40, loss: 3.537036900524981e-05
step: 50, loss: 0.0018851141212508082
step: 60, loss: 0.000488787773065269
step: 70, loss: 4.12580520787742e-05
step: 80, loss: 0.00010066706454381347
step: 90, loss: 0.0014269277453422546
step: 100, loss: 6.674870382994413e-05
step: 110, loss: 3.897371425409801e-05
step: 120, loss: 2.382256388955284e-05
step: 130, loss: 5.039535972173326e-05
step: 140, loss: 3.624188684625551e-05
step: 150, loss: 0.00023651207447983325
step: 160, loss: 4.3973512219963595e-05
step: 170, loss: 0.0043534450232982635
step: 180, loss: 8.805078687146306e-05
step: 190, loss: 0.00010353376273997128
step: 200, loss: 4.8717905883677304e-05
step: 210, loss: 3.088510493398644e-05
step: 220, loss: 8.750405686441809e-05
step: 230, loss: 7.314908725675195e-05
step: 240, loss: 0.0017805384704843163
step: 250, loss: 3.0602044716943055e-05
step: 260, loss: 2.878781560866628e-05

```

```

step: 270, loss: 3.253898466937244e-05
step: 280, loss: 3.998144165962003e-05
step: 290, loss: 3.1428164220415056e-05
step: 300, loss: 4.640278348233551e-05
step: 310, loss: 0.000928649737033993
step: 320, loss: 0.00127583765424788
step: 330, loss: 2.6582916689221747e-05
step: 340, loss: 4.986633939552121e-05
step: 350, loss: 3.449854557402432e-05
step: 360, loss: 3.827160981018096e-05
step: 370, loss: 4.46114172518719e-05
step: 380, loss: 0.00013176097127143294
step: 390, loss: 8.799511124379933e-05
step: 400, loss: 2.383365972491447e-05
step: 410, loss: 3.158548861392774e-05
step: 420, loss: 0.00018270053260494024
step: 430, loss: 0.00010483295045560226
step: 440, loss: 9.265793778467923e-05
step: 450, loss: 3.6498568078968674e-05
step: 460, loss: 0.0007317697745747864
step: 470, loss: 1.8890643332269974e-05
step: 480, loss: 1.9661627447931096e-05
step: 490, loss: 0.00012804257858078927
step: 500, loss: 1.5496982086915523e-05
step: 510, loss: 0.00012904613686259836
step: 520, loss: 0.00838453695178032
step: 530, loss: 2.236610271211248e-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=0.9060402684563759, f1=0.8977650974797908,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001879075716715306
step: 10, loss: 5.0281829317100346e-05
step: 20, loss: 0.00018034572713077068
step: 30, loss: 7.383129559457302e-05
step: 40, loss: 3.67821121471934e-05
step: 50, loss: 6.984344508964568e-05
step: 60, loss: 3.610967178246938e-05
step: 70, loss: 4.2806648707482964e-05
step: 80, loss: 0.00026506639551371336
step: 90, loss: 3.2892858143895864e-05
step: 100, loss: 0.00018273040768690407
step: 110, loss: 2.2444393835030496e-05
step: 120, loss: 3.9563336031278595e-05

```

```

step: 130, loss: 0.0007489445270039141
step: 140, loss: 0.00022063724463805556
step: 150, loss: 2.8355505492072552e-05
step: 160, loss: 4.899982741335407e-05
step: 170, loss: 2.6933128538075835e-05
step: 180, loss: 0.0005363999516703188
step: 190, loss: 0.0005667750956490636
step: 200, loss: 7.040639320621267e-05
step: 210, loss: 0.00010047260002465919
step: 220, loss: 3.4884353226516396e-05
step: 230, loss: 0.00014892089529894292
step: 240, loss: 0.0002466789446771145
step: 250, loss: 7.10891472408548e-05
step: 260, loss: 2.91866763291182e-05
step: 270, loss: 2.789009886328131e-05
step: 280, loss: 3.877202470903285e-05
step: 290, loss: 0.0006571727572008967
step: 300, loss: 2.593856152088847e-05
step: 310, loss: 0.00013692385982722044
step: 320, loss: 0.00010879508772632107
step: 330, loss: 0.049714308232069016
step: 340, loss: 8.991877984954044e-05
step: 350, loss: 4.0475479181623086e-05
step: 360, loss: 0.001993398880586028
step: 370, loss: 3.338076203363016e-05
step: 380, loss: 3.996131272288039e-05
step: 390, loss: 0.022813603281974792
step: 400, loss: 6.0853504692204297e-05
step: 410, loss: 2.301432323292829e-05
step: 420, loss: 2.3003054593573324e-05
step: 430, loss: 3.8726284401491284e-05
step: 440, loss: 0.002235251944512129
step: 450, loss: 4.388084926176816e-05
step: 460, loss: 3.495920464047231e-05
step: 470, loss: 0.0001639610854908824
step: 480, loss: 0.0001767695212038234
step: 490, loss: 2.6839670681511052e-05
step: 500, loss: 1.95984284800943e-05
step: 510, loss: 0.00023550588230136782
step: 520, loss: 1.7232961909030564e-05
step: 530, loss: 3.300800744909793e-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=0.9078080903104421, f1=0.9039812646370025,
best_f1=0.9110386585933862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6575518859317526e-05
step: 10, loss: 0.0002910557668656111
step: 20, loss: 2.290263000759296e-05
step: 30, loss: 0.0009789889445528388
step: 40, loss: 0.027383161708712578
step: 50, loss: 5.025218706578016e-05
step: 60, loss: 1.474078726459993e-05
step: 70, loss: 0.00021863561414647847
step: 80, loss: 0.0004995037452317774
step: 90, loss: 3.956828368245624e-05
step: 100, loss: 0.002054460346698761
step: 110, loss: 0.0016024120850488544
step: 120, loss: 0.0013309863861650229
step: 130, loss: 0.018934620544314384
step: 140, loss: 2.9289471058291383e-05
step: 150, loss: 3.8859929190948606e-05
step: 160, loss: 9.664092067396268e-05
step: 170, loss: 2.5294018996646628e-05
step: 180, loss: 2.1092115275678225e-05
step: 190, loss: 5.321435310179368e-05
step: 200, loss: 0.00030000830884091556
step: 210, loss: 0.0002753364678937942
step: 220, loss: 2.1557694708462805e-05
step: 230, loss: 5.238388985162601e-05
step: 240, loss: 2.246674739581067e-05
step: 250, loss: 2.8046590159647167e-05
step: 260, loss: 4.250480196787976e-05
step: 270, loss: 3.0017410608706996e-05
step: 280, loss: 0.0003507058136165142
step: 290, loss: 2.1013926016166806e-05
step: 300, loss: 3.038151771761477e-05
step: 310, loss: 0.0005897688679397106
step: 320, loss: 0.0009929003426805139
step: 330, loss: 3.2874006137717515e-05
step: 340, loss: 2.5803830794757232e-05
step: 350, loss: 1.728483766783029e-05
step: 360, loss: 4.7382767661474645e-05
step: 370, loss: 1.6368727301596664e-05
step: 380, loss: 2.626269997563213e-05
step: 390, loss: 8.167007763404399e-05
step: 400, loss: 4.759906732942909e-05
step: 410, loss: 2.571870390966069e-05
step: 420, loss: 4.1607534512877464e-05
step: 430, loss: 0.0007693520747125149
step: 440, loss: 2.533130464144051e-05
step: 450, loss: 2.7040379791287705e-05
step: 460, loss: 3.180884959874675e-05
```



```

step: 470, loss: 0.000394217116991058
step: 480, loss: 1.8115812054020353e-05
step: 490, loss: 2.545425741118379e-05
step: 500, loss: 0.004938072524964809
step: 510, loss: 3.670740261441097e-05
step: 520, loss: 2.0857462004642002e-05
step: 530, loss: 2.249657154607121e-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=0.9075391180654339, f1=0.9035916824196598,
best_f1=0.9110386585933862
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:22, 254.21it/s]
load_f1 = 0.9147141518275539
real_f1 = 0.9094287041337669
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:16, 265.83it/s]

2.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5426797866821289
step: 10, loss: 0.3923998475074768
step: 20, loss: 0.3950772285461426
step: 30, loss: 0.30407851934432983
step: 40, loss: 0.15551745891571045
step: 50, loss: 0.4896179437637329
step: 60, loss: 0.30525586009025574
step: 70, loss: 0.2600313425064087
step: 80, loss: 0.22365477681159973
step: 90, loss: 0.48140695691108704
step: 100, loss: 0.5283660292625427
step: 110, loss: 0.3026462495326996
step: 120, loss: 0.30361634492874146
step: 130, loss: 0.4376980662345886
step: 140, loss: 0.22929660975933075
step: 150, loss: 0.36233311891555786
step: 160, loss: 0.2941146790981293
step: 170, loss: 0.355407178401947
step: 180, loss: 0.21316996216773987
step: 190, loss: 0.2522808611392975
step: 200, loss: 0.349767804145813
step: 210, loss: 0.25462666153907776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.41234567901234576, f1=0.42051282051282046,
best_f1=0.42051282051282046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17819936573505402
step: 10, loss: 0.4076669216156006
step: 20, loss: 0.2884460389614105
step: 30, loss: 0.3635062873363495
step: 40, loss: 0.22958897054195404
step: 50, loss: 0.15024986863136292
step: 60, loss: 0.47098755836486816
step: 70, loss: 0.153959259390831
step: 80, loss: 0.2959614098072052
step: 90, loss: 0.11744856834411621
step: 100, loss: 0.0408531092107296
step: 110, loss: 0.1835208237171173
step: 120, loss: 0.1610432118177414
step: 130, loss: 0.055518172681331635
step: 140, loss: 0.22950586676597595
step: 150, loss: 0.24987098574638367
step: 160, loss: 0.2344844937324524
step: 170, loss: 0.16472913324832916
step: 180, loss: 0.2561550736427307
step: 190, loss: 0.25989991426467896
step: 200, loss: 0.058524757623672485
step: 210, loss: 0.22734874486923218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4333333333333333, f1=0.49218750000000006,
best_f1=0.49218750000000006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08735973387956619
step: 10, loss: 0.34220045804977417
step: 20, loss: 0.19989433884620667
step: 30, loss: 0.21439698338508606
step: 40, loss: 0.08840638399124146
step: 50, loss: 0.10576336830854416
```

```

step: 60, loss: 0.12677665054798126
step: 70, loss: 0.33543044328689575
step: 80, loss: 0.1408807933330536
step: 90, loss: 0.12940475344657898
step: 100, loss: 0.13640831410884857
step: 110, loss: 0.3022952973842621
step: 120, loss: 0.1825467199087143
step: 130, loss: 0.14052678644657135
step: 140, loss: 0.13494467735290527
step: 150, loss: 0.245692640542984
step: 160, loss: 0.04371216148138046
step: 170, loss: 0.1540675014257431
step: 180, loss: 0.15874351561069489
step: 190, loss: 0.3344349265098572
step: 200, loss: 0.15391311049461365
step: 210, loss: 0.10943134129047394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.481675392670157, f1=0.5501730103806229,
best_f1=0.5501730103806229
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21451805531978607
step: 10, loss: 0.21123136579990387
step: 20, loss: 0.07878239452838898
step: 30, loss: 0.15986643731594086
step: 40, loss: 0.052230652421712875
step: 50, loss: 0.21481625735759735
step: 60, loss: 0.23839086294174194
step: 70, loss: 0.20703968405723572
step: 80, loss: 0.11679672449827194
step: 90, loss: 0.04422552138566971
step: 100, loss: 0.3668478727340698
step: 110, loss: 0.17491081357002258
step: 120, loss: 0.18916144967079163
step: 130, loss: 0.10921388119459152
step: 140, loss: 0.24669653177261353
step: 150, loss: 0.06275449693202972
step: 160, loss: 0.06088748201727867
step: 170, loss: 0.20628094673156738
step: 180, loss: 0.23743271827697754
step: 190, loss: 0.1325361132621765

```

```

step: 200, loss: 0.24484850466251373
step: 210, loss: 0.282761812210083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.49612403100775193, f1=0.505091649694501,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14892442524433136
step: 10, loss: 0.06270725280046463
step: 20, loss: 0.33654868602752686
step: 30, loss: 0.03935299441218376
step: 40, loss: 0.06843771785497665
step: 50, loss: 0.14941994845867157
step: 60, loss: 0.08641224354505539
step: 70, loss: 0.03504486009478569
step: 80, loss: 0.03760456293821335
step: 90, loss: 0.07081558555364609
step: 100, loss: 0.021763980388641357
step: 110, loss: 0.19513729214668274
step: 120, loss: 0.09140560030937195
step: 130, loss: 0.11524120718240738
step: 140, loss: 0.16212211549282074
step: 150, loss: 0.0978723019361496
step: 160, loss: 0.11911187320947647
step: 170, loss: 0.08491872996091843
step: 180, loss: 0.07323506474494934
step: 190, loss: 0.02328369952738285
step: 200, loss: 0.20881254971027374
step: 210, loss: 0.03794702887535095
/usr/local/lib/python3.7/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=0.4747081712062257, f1=0.5156250000000001,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07848545163869858
step: 10, loss: 0.027983125299215317
step: 20, loss: 0.07775142788887024
step: 30, loss: 0.004306306131184101

```

```

step: 40, loss: 0.05558066815137863
step: 50, loss: 0.008415297605097294
step: 60, loss: 0.033098042011260986
step: 70, loss: 0.00888659618794918
step: 80, loss: 0.008771886117756367
step: 90, loss: 0.20177745819091797
step: 100, loss: 0.010179208591580391
step: 110, loss: 0.06651108711957932
step: 120, loss: 0.04549430310726166
step: 130, loss: 0.07970115542411804
step: 140, loss: 0.042611442506313324
step: 150, loss: 0.051211901009082794
step: 160, loss: 0.004759099800139666
step: 170, loss: 0.17464500665664673
step: 180, loss: 0.1754651963710785
step: 190, loss: 0.14118723571300507
step: 200, loss: 0.05590846762061119
step: 210, loss: 0.03986458480358124
/usr/local/lib/python3.7/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=0.4729458917835671, f1=0.5000000000000001,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07782814651727676
step: 10, loss: 0.018723521381616592
step: 20, loss: 0.0351259745657444
step: 30, loss: 0.08481337130069733
step: 40, loss: 0.013196388259530067
step: 50, loss: 0.17918719351291656
step: 60, loss: 0.01157052256166935
step: 70, loss: 0.03734205290675163
step: 80, loss: 0.1152462437748909
step: 90, loss: 0.03200991451740265
step: 100, loss: 0.003628677921369672
step: 110, loss: 0.21404916048049927
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.10006233304738998
step: 130, loss: 0.08105587214231491
step: 140, loss: 0.053723566234111786
step: 150, loss: 0.04881145805120468
step: 160, loss: 0.015621690079569817
step: 170, loss: 0.001384107512421906
step: 180, loss: 0.038077834993600845
step: 190, loss: 0.00578369852155447
step: 200, loss: 0.016077658161520958

```

```

step: 210, loss: 0.05945159122347832
/usr/local/lib/python3.7/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=0.4714828897338402, f1=0.5265225933202357,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02164926566183567
step: 10, loss: 0.05822917819023132
step: 20, loss: 0.0025538827758282423
step: 30, loss: 0.00611715717241168
step: 40, loss: 0.09965914487838745
step: 50, loss: 0.004360053688287735
step: 60, loss: 0.005616145674139261
step: 70, loss: 0.02029598504304886
step: 80, loss: 0.035412829369306564
step: 90, loss: 0.03789268061518669
step: 100, loss: 0.111116923391819
step: 110, loss: 0.130131334066391
step: 120, loss: 0.02298450842499733
step: 130, loss: 0.03069094941020012
step: 140, loss: 0.06964189559221268
step: 150, loss: 0.08773119747638702
step: 160, loss: 0.3261536657810211
step: 170, loss: 0.005899264942854643
step: 180, loss: 0.06650515645742416
step: 190, loss: 0.009789623320102692
step: 200, loss: 0.0030400315299630165
step: 210, loss: 0.08760897070169449
/usr/local/lib/python3.7/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=0.48163265306122444, f1=0.5409836065573771,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039012789726257324
step: 10, loss: 0.017656484618782997
step: 20, loss: 0.006377456244081259
step: 30, loss: 0.010090194642543793
step: 40, loss: 0.1197633296251297
step: 50, loss: 0.04368444159626961
step: 60, loss: 0.08369807153940201
step: 70, loss: 0.02228277549147606
step: 80, loss: 0.0016394879203289747

```



```

step: 90, loss: 0.0016079063061624765
step: 100, loss: 0.009774786420166492
step: 110, loss: 0.004713958129286766
step: 120, loss: 0.0002979193232022226
step: 130, loss: 0.008427446708083153
step: 140, loss: 0.003709869459271431
step: 150, loss: 0.061958275735378265
step: 160, loss: 0.04424503073096275
step: 170, loss: 0.05959422513842583
step: 180, loss: 0.008964551612734795
step: 190, loss: 0.0008269580430351198
step: 200, loss: 0.06743938475847244
step: 210, loss: 0.0026386112440377474
/usr/local/lib/python3.7/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=0.4222222222222222, f1=0.48106904231625836,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08231578767299652
step: 10, loss: 0.027477331459522247
step: 20, loss: 0.004044689238071442
step: 30, loss: 0.0571649968624115
step: 40, loss: 0.005517024081200361
step: 50, loss: 0.013484789058566093
step: 60, loss: 0.001250320696271956
step: 70, loss: 0.015206036157906055
step: 80, loss: 0.01703428104519844
step: 90, loss: 0.005543222185224295
step: 100, loss: 0.04498358815908432
step: 110, loss: 0.0008644724148325622
step: 120, loss: 0.11407414078712463
step: 130, loss: 0.0170478206127882
step: 140, loss: 0.00036279112100601196
step: 150, loss: 0.012391861528158188
step: 160, loss: 0.0027786321006715298
step: 170, loss: 0.0013204109854996204
step: 180, loss: 0.021040627732872963
step: 190, loss: 0.03897082805633545
step: 200, loss: 0.0019116714829578996
step: 210, loss: 0.0392032153904438
/usr/local/lib/python3.7/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=0.46189376443418007, f1=0.5046296296296297,
best_f1=0.505091649694501

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005312936380505562
step: 10, loss: 0.04680482670664787
step: 20, loss: 0.001249426626600325
step: 30, loss: 0.04293788596987724
step: 40, loss: 0.006174099165946245
step: 50, loss: 0.021346716210246086
step: 60, loss: 0.0014023862313479185
step: 70, loss: 0.0016508324770256877
step: 80, loss: 0.015383332967758179
step: 90, loss: 0.014797812327742577
step: 100, loss: 0.021555010229349136
step: 110, loss: 0.008500649593770504
step: 120, loss: 0.029028022661805153
step: 130, loss: 0.0022770778741687536
step: 140, loss: 0.0013822013279423118
step: 150, loss: 0.00012411335774231702
step: 160, loss: 0.003914321307092905
step: 170, loss: 0.001786491833627224
step: 180, loss: 0.018365759402513504
step: 190, loss: 0.0015768560115247965
step: 200, loss: 0.024192288517951965
step: 210, loss: 0.0022031180560588837
/usr/local/lib/python3.7/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=0.45585215605749485, f1=0.5220883534136546,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001655207364819944
step: 10, loss: 0.007558510173112154
step: 20, loss: 0.0564366951584816
step: 30, loss: 0.0035009668208658695
step: 40, loss: 0.01663283072412014
step: 50, loss: 0.00024393032072111964
step: 60, loss: 0.005604021716862917
step: 70, loss: 0.01993144303560257
step: 80, loss: 0.0002491529448889196
step: 90, loss: 0.029563257470726967
step: 100, loss: 0.0023146390449255705
step: 110, loss: 0.00522651057690382
step: 120, loss: 0.0008352509466931224
step: 130, loss: 0.0008349798736162484
step: 140, loss: 0.0002833227044902742

```

```

step: 150, loss: 0.002494232030585408
step: 160, loss: 0.0003075031563639641
step: 170, loss: 0.00018096662824973464
step: 180, loss: 0.01673831418156624
step: 190, loss: 0.0011949193431064487
step: 200, loss: 0.0002586132031865418
step: 210, loss: 0.0024699338246136904
/usr/local/lib/python3.7/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=0.44052863436123346, f1=0.5161290322580645,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015394758433103561
step: 10, loss: 0.0002151781227439642
step: 20, loss: 0.0002111108769895509
step: 30, loss: 0.001651874976232648
step: 40, loss: 0.00017491051403339952
step: 50, loss: 0.0008790877182036638
step: 60, loss: 0.0005668403464369476
step: 70, loss: 0.002955404343083501
step: 80, loss: 0.0013564785476773977
step: 90, loss: 0.00023272463295143098
step: 100, loss: 0.0001713066449156031
step: 110, loss: 0.00022641710529569536
step: 120, loss: 0.022184329107403755
step: 130, loss: 0.00020380491332616657
step: 140, loss: 0.00014093164645601064
step: 150, loss: 0.0012819968396797776
step: 160, loss: 0.0005831727758049965
step: 170, loss: 0.018146540969610214
step: 180, loss: 0.0013225963339209557
step: 190, loss: 0.0010838336311280727
step: 200, loss: 0.008309069089591503
step: 210, loss: 0.0007335490663535893
/usr/local/lib/python3.7/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=0.43280182232346237, f1=0.49777777777777776,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007095240871421993
step: 10, loss: 0.04485756903886795
step: 20, loss: 0.0013171248137950897

```

```

step: 30, loss: 0.004214869346469641
step: 40, loss: 0.00014857131463941187
step: 50, loss: 0.0013864220818504691
step: 60, loss: 0.007738546002656221
step: 70, loss: 0.0006319441017694771
step: 80, loss: 0.0012670642463490367
step: 90, loss: 0.0001706987532088533
step: 100, loss: 0.0005300985067151487
step: 110, loss: 0.0001973291509784758
step: 120, loss: 0.0015170825645327568
step: 130, loss: 0.0009758021333254874
step: 140, loss: 0.01429979782551527
step: 150, loss: 0.0002626234490890056
step: 160, loss: 0.00018237385665997863
step: 170, loss: 0.00021152796398382634
step: 180, loss: 0.09181705117225647
step: 190, loss: 0.0007036340539343655
step: 200, loss: 0.0008369316346943378
step: 210, loss: 0.0060887872241437435
/usr/local/lib/python3.7/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=0.4403669724770642, f1=0.49549549549549554,
best_f1=0.505091649694501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026720267487689853
step: 10, loss: 0.00023107031302060932
step: 20, loss: 0.0004475457826629281
step: 30, loss: 0.00795670971274376
step: 40, loss: 0.0005005674320273101
step: 50, loss: 0.00023686785425525159
step: 60, loss: 0.05950000137090683
step: 70, loss: 0.003180940169841051
step: 80, loss: 0.00027246124227531254
step: 90, loss: 0.00287640537135303
step: 100, loss: 0.00019920972408726811
step: 110, loss: 0.00013743434101343155
step: 120, loss: 0.00021299743093550205
step: 130, loss: 0.00020418014901224524
step: 140, loss: 0.00010988357098540291
step: 150, loss: 0.00010398389713373035
step: 160, loss: 0.0022685816511511803
step: 170, loss: 0.0006768582388758659
step: 180, loss: 0.0010077544720843434
step: 190, loss: 0.0001915024040499702
step: 200, loss: 0.018787279725074768

```

```

step: 210, loss: 0.0014967890456318855
/usr/local/lib/python3.7/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=0.4409090909090909, f1=0.5077262693156732,
best_f1=0.505091649694501
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:05, 433.88it/s]
load_f1 = 0.48565965583174
real_f1 = 0.4980544747081712
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 265.84it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5551677346229553
step: 10, loss: 0.38359037041664124
step: 20, loss: 0.29695695638656616
step: 30, loss: 0.44376951456069946
step: 40, loss: 0.42356643080711365
step: 50, loss: 0.3159785270690918
step: 60, loss: 0.3166933059692383
step: 70, loss: 0.3018368184566498
step: 80, loss: 0.23721438646316528
step: 90, loss: 0.29733479022979736
step: 100, loss: 0.27518168091773987
step: 110, loss: 0.5228491425514221
step: 120, loss: 0.21038775146007538
step: 130, loss: 0.16287943720817566
step: 140, loss: 0.09186311066150665
step: 150, loss: 0.13551662862300873
step: 160, loss: 0.15972062945365906
step: 170, loss: 0.25266143679618835
step: 180, loss: 0.0436091274023056
step: 190, loss: 0.3854607045650482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.640552995391705, f1=0.6428571428571428,
best_f1=0.6428571428571428

```

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

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

```
step: 0, loss: 0.2221403270959854  
step: 10, loss: 0.10855066776275635  
step: 20, loss: 0.16377851366996765  
step: 30, loss: 0.1345938742160797  
step: 40, loss: 0.14333857595920563  
step: 50, loss: 0.07474817335605621  
step: 60, loss: 0.26793864369392395  
step: 70, loss: 0.22370386123657227  
step: 80, loss: 0.1473010778427124  
step: 90, loss: 0.11567679047584534  
step: 100, loss: 0.022125648334622383  
step: 110, loss: 0.2648455500602722  
step: 120, loss: 0.2390345185995102  
step: 130, loss: 0.13395695388317108  
step: 140, loss: 0.08726227283477783  
step: 150, loss: 0.1918906271457672  
step: 160, loss: 0.06935739517211914  
step: 170, loss: 0.18284252285957336  
step: 180, loss: 0.13743746280670166  
step: 190, loss: 0.05020506680011749
```

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

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

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

```
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=0.7354497354497355, f1=0.7098445595854923,  
best_f1=0.7098445595854923
```

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

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

```
step: 0, loss: 0.04093864932656288  
step: 10, loss: 0.24724172055721283  
step: 20, loss: 0.21029512584209442  
step: 30, loss: 0.2690320312976837  
step: 40, loss: 0.11733613908290863  
step: 50, loss: 0.2099025845527649  
step: 60, loss: 0.014912514947354794  
step: 70, loss: 0.11107493191957474  
step: 80, loss: 0.02799360081553459  
step: 90, loss: 0.089402936398983  
step: 100, loss: 0.10025516897439957  
step: 110, loss: 0.009232338517904282  
step: 120, loss: 0.08510759472846985
```



```

step: 130, loss: 0.014243008568882942
step: 140, loss: 0.05518047511577606
step: 150, loss: 0.11189482361078262
step: 160, loss: 0.02781740017235279
step: 170, loss: 0.2009870558977127
step: 180, loss: 0.018007805570960045
step: 190, loss: 0.11911854147911072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7857142857142857, f1=0.7675675675675676,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05300365760922432
step: 10, loss: 0.15596450865268707
step: 20, loss: 0.07253675162792206
step: 30, loss: 0.09168107062578201
step: 40, loss: 0.08551938086748123
step: 50, loss: 0.043563008308410645
step: 60, loss: 0.1160314679145813
step: 70, loss: 0.07791784405708313
step: 80, loss: 0.13772980868816376
step: 90, loss: 0.05422919988632202
step: 100, loss: 0.1084827110171318
step: 110, loss: 0.047379299998283386
step: 120, loss: 0.010984751395881176
step: 130, loss: 0.04561397060751915
step: 140, loss: 0.01655276119709015
step: 150, loss: 0.01694018766283989
step: 160, loss: 0.017888305708765984
step: 170, loss: 0.17542608082294464
step: 180, loss: 0.022888291627168655
step: 190, loss: 0.09803203493356705
/usr/local/lib/python3.7/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=0.7712765957446809, f1=0.761904761904762,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044922854751348495

```

```

step: 10, loss: 0.015385533683001995
step: 20, loss: 0.040202200412750244
step: 30, loss: 0.011408987455070019
step: 40, loss: 0.02331758663058281
step: 50, loss: 0.19494660198688507
step: 60, loss: 0.08943741768598557
step: 70, loss: 0.009465666487812996
step: 80, loss: 0.014021703042089939
step: 90, loss: 0.009402510710060596
step: 100, loss: 0.0025155607145279646
step: 110, loss: 0.004278331529349089
step: 120, loss: 0.001849177642725408
step: 130, loss: 0.011202544905245304
step: 140, loss: 0.37154102325439453
step: 150, loss: 0.09226171672344208
step: 160, loss: 0.14680345356464386
step: 170, loss: 0.07251042872667313
step: 180, loss: 0.12537366151809692
step: 190, loss: 0.2049197107553482
/usr/local/lib/python3.7/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=0.7232876712328767, f1=0.7353760445682451,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06896695494651794
step: 10, loss: 0.03170429542660713
step: 20, loss: 0.0036820622626692057
step: 30, loss: 0.10992524772882462
step: 40, loss: 0.005419062916189432
step: 50, loss: 0.02198847010731697
step: 60, loss: 0.0019710955675691366
step: 70, loss: 0.015240645967423916
step: 80, loss: 0.011301778256893158
step: 90, loss: 0.005617510061711073
step: 100, loss: 0.0005458440282382071
step: 110, loss: 0.005616160575300455
step: 120, loss: 0.19570060074329376
step: 130, loss: 0.05123564973473549
step: 140, loss: 0.005030622705817223
step: 150, loss: 0.005234931129962206
step: 160, loss: 0.021166659891605377
step: 170, loss: 0.05919670686125755
step: 180, loss: 0.075951486825943
step: 190, loss: 0.11114110797643661
/usr/local/lib/python3.7/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=0.7642276422764227, f1=0.7527472527472527,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009750358760356903
step: 10, loss: 0.005962917115539312
step: 20, loss: 0.021210290491580963
step: 30, loss: 0.014474866911768913
step: 40, loss: 0.04917248710989952
step: 50, loss: 0.07992348819971085
step: 60, loss: 0.0020647633355110884
step: 70, loss: 0.0020983752328902483
step: 80, loss: 0.05156882852315903
step: 90, loss: 0.0010492000728845596
step: 100, loss: 0.0014665351482108235
step: 110, loss: 0.003760933643206954
step: 120, loss: 0.026601428166031837
step: 130, loss: 0.001510730478912592
step: 140, loss: 0.024633260443806648
step: 150, loss: 0.1977282464504242
step: 160, loss: 0.10265371948480606
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.022301912307739258
step: 180, loss: 0.0023347476962953806
step: 190, loss: 0.0016348259523510933
/usr/local/lib/python3.7/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=0.752808988764045, f1=0.7272727272727273,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007292592898011208
step: 10, loss: 0.005497207399457693
step: 20, loss: 0.0009310589521192014
step: 30, loss: 0.0022910558618605137
step: 40, loss: 0.0005309623084031045
step: 50, loss: 0.000429505598731339
step: 60, loss: 0.0009296414791606367
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.0017058803932741284
step: 80, loss: 0.00031516465242020786
step: 90, loss: 0.003367889206856489
step: 100, loss: 0.042944904416799545

```

```

step: 110, loss: 0.0007353078108280897
step: 120, loss: 0.0020421352237462997
step: 130, loss: 0.03720752149820328
step: 140, loss: 0.010813808999955654
step: 150, loss: 0.05591299384832382
step: 160, loss: 0.022632071748375893
step: 170, loss: 0.015405556187033653
step: 180, loss: 0.07490944862365723
step: 190, loss: 0.0021211798302829266
/usr/local/lib/python3.7/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=0.7828418230563003, f1=0.7439353099730459,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03373706713318825
step: 10, loss: 0.02128874510526657
step: 20, loss: 0.08927545696496964
step: 30, loss: 0.009721996262669563
step: 40, loss: 0.0005860962555743754
step: 50, loss: 0.00043523142812773585
step: 60, loss: 0.00030168838566169143
step: 70, loss: 0.034403447061777115
step: 80, loss: 0.00036271955468691885
step: 90, loss: 0.012175235897302628
step: 100, loss: 0.03080364689230919
step: 110, loss: 0.0055426014587283134
step: 120, loss: 0.005022811237722635
step: 130, loss: 0.016906777396798134
step: 140, loss: 0.022617755457758904
step: 150, loss: 0.021280279383063316
step: 160, loss: 0.0006786949234083295
step: 170, loss: 0.0005030982429161668
step: 180, loss: 0.03570754826068878
step: 190, loss: 0.09027548879384995
/usr/local/lib/python3.7/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=0.7828418230563003, f1=0.7634408602150539,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006290496094152331
step: 10, loss: 0.0011984106386080384
step: 20, loss: 0.0007069961866363883

```

```

step: 30, loss: 0.00043042234028689563
step: 40, loss: 0.0015255825128406286
step: 50, loss: 0.0008250540122389793
step: 60, loss: 0.0061111655086278915
step: 70, loss: 0.019111553207039833
step: 80, loss: 0.01853160932660103
step: 90, loss: 0.0004750483203679323
step: 100, loss: 0.002086711348965764
step: 110, loss: 0.0003779087564907968
step: 120, loss: 0.1554163098335266
step: 130, loss: 0.0003171016869600862
step: 140, loss: 0.0007106553530320525
step: 150, loss: 0.012983487918972969
step: 160, loss: 0.007816221565008163
step: 170, loss: 0.00028649321757256985
step: 180, loss: 0.0021060339640825987
step: 190, loss: 0.00033806206192821264
/usr/local/lib/python3.7/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=0.7726027397260273, f1=0.7380281690140845,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033828968298621476
step: 10, loss: 0.004776752088218927
step: 20, loss: 0.0028422437608242035
step: 30, loss: 0.007857165299355984
step: 40, loss: 0.0006839917041361332
step: 50, loss: 0.0033321750815957785
step: 60, loss: 0.00022052160056773573
step: 70, loss: 0.00016133137978613377
step: 80, loss: 0.00018312076281290501
step: 90, loss: 0.00019277801038697362
step: 100, loss: 0.00028952083084732294
step: 110, loss: 0.00028869142988696694
step: 120, loss: 0.00020423783280421048
step: 130, loss: 0.0003049120132345706
step: 140, loss: 0.0003075219865422696
step: 150, loss: 0.0002282688074046746
step: 160, loss: 0.00023255647101905197
step: 170, loss: 0.00018740688392426819
step: 180, loss: 0.0012648575939238071
step: 190, loss: 0.0002768216363620013
/usr/local/lib/python3.7/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=0.7774647887323942, f1=0.7428571428571429,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.347918628714979e-05
step: 10, loss: 0.00020531169138848782
step: 20, loss: 0.00012918970605824143
step: 30, loss: 0.0003385838062968105
step: 40, loss: 0.0004447795799933374
step: 50, loss: 0.0002791833830997348
step: 60, loss: 0.00011336930037941784
step: 70, loss: 9.98809773591347e-05
step: 80, loss: 0.0001466138637624681
step: 90, loss: 0.0001639363035792485
step: 100, loss: 0.0002112120419042185
step: 110, loss: 0.0001577071234351024
step: 120, loss: 0.00019264592265244573
step: 130, loss: 0.0025633724872022867
step: 140, loss: 0.007478238083422184
step: 150, loss: 0.00019488493853714317
step: 160, loss: 0.00022421858739107847
step: 170, loss: 0.000417455768911168
step: 180, loss: 0.0001409105898346752
step: 190, loss: 0.00011888190783793107
/usr/local/lib/python3.7/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=0.7811634349030471, f1=0.7457627118644068,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000430869753472507
step: 10, loss: 0.00017145044694188982
step: 20, loss: 0.0010004742071032524
step: 30, loss: 0.0008660274324938655
step: 40, loss: 0.0001363895135000348
step: 50, loss: 0.00019430987595114857
step: 60, loss: 0.00034322438295930624
step: 70, loss: 0.002529968973249197
step: 80, loss: 0.00016289713676087558
step: 90, loss: 0.00043444643961265683
step: 100, loss: 0.00019439475727267563
step: 110, loss: 9.54652568907477e-05
step: 120, loss: 0.0002131026703864336
step: 130, loss: 0.00010361584281781688
step: 140, loss: 9.010804205900058e-05

```

```

step: 150, loss: 0.00011410920706111938
step: 160, loss: 0.0002313855366082862
step: 170, loss: 0.000273957964964211
step: 180, loss: 0.00011210668890271336
step: 190, loss: 0.00037149229319766164
/usr/local/lib/python3.7/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=0.7810026385224274, f1=0.7465940054495912,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000575369456782937
step: 10, loss: 0.0002111146313836798
step: 20, loss: 0.00034412628156132996
step: 30, loss: 9.29335001274012e-05
step: 40, loss: 0.00808777566999197
step: 50, loss: 0.0010430009569972754
step: 60, loss: 0.00016054423758760095
step: 70, loss: 0.0001207397217513062
step: 80, loss: 0.00037134604644961655
step: 90, loss: 0.00046604813542217016
step: 100, loss: 0.0034909467212855816
step: 110, loss: 0.00018146811635233462
step: 120, loss: 0.0004372734110802412
step: 130, loss: 0.0001560931996209547
step: 140, loss: 0.000295478937914595
step: 150, loss: 0.0001415312581229955
step: 160, loss: 0.00032256124541163445
step: 170, loss: 0.0001936815824592486
step: 180, loss: 0.00012146853259764612
step: 190, loss: 0.0003203271480742842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.786096256684492, f1=0.7431693989071039,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008161081932485104
step: 10, loss: 0.00021856525563634932
step: 20, loss: 0.00013450018013827503

```

```

step: 30, loss: 0.0012072203680872917
step: 40, loss: 0.00018584023928269744
step: 50, loss: 0.00028290157206356525
step: 60, loss: 0.00010189716704189777
step: 70, loss: 0.0003195599711034447
step: 80, loss: 0.0006400900892913342
step: 90, loss: 0.00017574543016962707
step: 100, loss: 0.0041572535410523415
step: 110, loss: 0.00023803333169780672
step: 120, loss: 0.00010936654143733904
step: 130, loss: 0.00014850379375275224
step: 140, loss: 0.0005771336727775633
step: 150, loss: 0.00039717109757475555
step: 160, loss: 0.00019458284077700227
step: 170, loss: 0.00013581950042862445
step: 180, loss: 0.0003800082195084542
step: 190, loss: 0.00014857295900583267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7925531914893618, f1=0.7465940054495912,
best_f1=0.7465940054495912
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:08, 231.38it/s]
load_f1 = 0.7890410958904108
real_f1 = 0.7713498622589531
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 250.03it/s]

```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

```

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

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

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

```

```
--checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result

in PyTorch skipping the first value of the learning rate schedule. See more

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

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

```
step: 0, loss: 0.6136382818222046
step: 10, loss: 0.3752635717391968
step: 20, loss: 0.3068855106830597
step: 30, loss: 0.37660157680511475
step: 40, loss: 0.27966034412384033
step: 50, loss: 0.23037798702716827
step: 60, loss: 0.30434393882751465
step: 70, loss: 0.3531155586242676
step: 80, loss: 0.32512933015823364
```

```

step: 90, loss: 0.15566490590572357
step: 100, loss: 0.23830054700374603
step: 110, loss: 0.20853351056575775
step: 120, loss: 0.14859884977340698
step: 130, loss: 0.0649643987417221
step: 140, loss: 0.2281384915113449
step: 150, loss: 0.26947423815727234
step: 160, loss: 0.13255757093429565
step: 170, loss: 0.3162281811237335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7281553398058251, f1=0.6714628297362111,
best_f1=0.6714628297362111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036242179572582245
step: 10, loss: 0.11107733845710754
step: 20, loss: 0.1057780459523201
step: 30, loss: 0.21576346457004547
step: 40, loss: 0.12152107059955597
step: 50, loss: 0.09643299877643585
step: 60, loss: 0.41632628440856934
step: 70, loss: 0.09719308465719223
step: 80, loss: 0.04123036190867424
step: 90, loss: 0.11312873661518097
step: 100, loss: 0.1294376403093338
step: 110, loss: 0.08647368848323822
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.10153007507324219
step: 130, loss: 0.10523241013288498
step: 140, loss: 0.22545203566551208
step: 150, loss: 0.21166828274726868
step: 160, loss: 0.07018417119979858
step: 170, loss: 0.06087721139192581
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7524271844660194, f1=0.7294685990338163,

```

```

best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13905741274356842
step: 10, loss: 0.05953509733080864
step: 20, loss: 0.05399194359779358
step: 30, loss: 0.07503440231084824
step: 40, loss: 0.08895828574895859
step: 50, loss: 0.08294324576854706
step: 60, loss: 0.11327610909938812
step: 70, loss: 0.09562437236309052
step: 80, loss: 0.08622594177722931
step: 90, loss: 0.1608278453350067
step: 100, loss: 0.025483854115009308
step: 110, loss: 0.20195835828781128
step: 120, loss: 0.07559292763471603
step: 130, loss: 0.11598372459411621
step: 140, loss: 0.13114763796329498
step: 150, loss: 0.014833083376288414
step: 160, loss: 0.06634525209665298
step: 170, loss: 0.05638795346021652
/usr/local/lib/python3.7/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=0.745945945945946, f1=0.7433155080213903,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11892732977867126
step: 10, loss: 0.03574776649475098
step: 20, loss: 0.009932862594723701
step: 30, loss: 0.04237498342990875
step: 40, loss: 0.001962742069736123
step: 50, loss: 0.10910174995660782
step: 60, loss: 0.07557125389575958
step: 70, loss: 0.06339041143655777
step: 80, loss: 0.12360205501317978
step: 90, loss: 0.07473418861627579
step: 100, loss: 0.5073763728141785
step: 110, loss: 0.06123225390911102
step: 120, loss: 0.009880597703158855
step: 130, loss: 0.1798718124628067
step: 140, loss: 0.04904408007860184
step: 150, loss: 0.02201320417225361
step: 160, loss: 0.13331389427185059
step: 170, loss: 0.01399607490748167

```

```

/usr/local/lib/python3.7/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=0.730478589420655, f1=0.6947890818858561,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018580660223960876
step: 10, loss: 0.03721736744046211
step: 20, loss: 0.010900291614234447
step: 30, loss: 0.040283091366291046
step: 40, loss: 0.015274765901267529
step: 50, loss: 0.040939319878816605
step: 60, loss: 0.01577875204384327
step: 70, loss: 0.23692308366298676
step: 80, loss: 0.06684126704931259
step: 90, loss: 0.10213816165924072
step: 100, loss: 0.05511562153697014
step: 110, loss: 0.05012292042374611
step: 120, loss: 0.01960013434290886
step: 130, loss: 0.06582170724868774
step: 140, loss: 0.07388903200626373
step: 150, loss: 0.05606750398874283
step: 160, loss: 0.03125794231891632
step: 170, loss: 0.09486236423254013
/usr/local/lib/python3.7/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=0.7118644067796611, f1=0.6848072562358277,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011037887074053288
step: 10, loss: 0.020505085587501526
step: 20, loss: 0.04473388195037842
step: 30, loss: 0.025697069242596626
step: 40, loss: 0.01657099276781082
step: 50, loss: 0.11525344103574753
step: 60, loss: 0.06176977604627609
step: 70, loss: 0.054364610463380814
step: 80, loss: 0.03600839897990227
step: 90, loss: 0.01877853460609913
step: 100, loss: 0.003729012329131365
step: 110, loss: 0.005265751387923956
step: 120, loss: 0.04053587466478348
step: 130, loss: 0.002705428982153535

```

```

step: 140, loss: 0.04675706475973129
step: 150, loss: 0.11883704364299774
step: 160, loss: 0.07123296707868576
step: 170, loss: 0.00922603253275156
/usr/local/lib/python3.7/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=0.7259615384615385, f1=0.7058823529411764,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029241524171084166
step: 10, loss: 0.01683959923684597
step: 20, loss: 0.012110122479498386
step: 30, loss: 0.05361044034361839
step: 40, loss: 0.025731857866048813
step: 50, loss: 0.03576033189892769
step: 60, loss: 0.008382371626794338
step: 70, loss: 0.013215750455856323
step: 80, loss: 0.0023470469750463963
step: 90, loss: 0.00043265620479360223
step: 100, loss: 0.006412746850401163
step: 110, loss: 0.05877241864800453
step: 120, loss: 0.05057545378804207
step: 130, loss: 0.0891028419137001
step: 140, loss: 0.006285690236836672
step: 150, loss: 0.023829080164432526
step: 160, loss: 0.013160414062440395
step: 170, loss: 0.049793507903814316
/usr/local/lib/python3.7/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=0.7210526315789473, f1=0.7376623376623377,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003632546868175268
step: 10, loss: 0.004323170054703951
step: 20, loss: 0.001580391894094646
step: 30, loss: 0.034204501658678055
step: 40, loss: 0.0003508379450067878
step: 50, loss: 0.00268789054825902
step: 60, loss: 0.002802557311952114
step: 70, loss: 0.02467915415763855
step: 80, loss: 0.010346595197916031
step: 90, loss: 0.006067960523068905

```

```

step: 100, loss: 0.0357830747961998
step: 110, loss: 0.027352385222911835
step: 120, loss: 0.0012073492398485541
step: 130, loss: 0.001068966812454164
step: 140, loss: 0.0006005583563819528
step: 150, loss: 0.0018562785116955638
step: 160, loss: 0.010641630738973618
step: 170, loss: 0.020235218107700348
/usr/local/lib/python3.7/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=0.7070707070707071, f1=0.6904761904761905,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023444803431630135
step: 10, loss: 0.01692323572933674
step: 20, loss: 0.03893419727683067
step: 30, loss: 0.005962858907878399
step: 40, loss: 0.016843808814883232
step: 50, loss: 0.001180503168143332
step: 60, loss: 0.002662557177245617
step: 70, loss: 0.03397315368056297
step: 80, loss: 0.00928837526589632
step: 90, loss: 0.013962341472506523
step: 100, loss: 0.007804979104548693
step: 110, loss: 0.13527537882328033
step: 120, loss: 0.003185845213010907
step: 130, loss: 0.01476262602955103
step: 140, loss: 0.001722769346088171
step: 150, loss: 0.013433174230158329
step: 160, loss: 0.09735722094774246
step: 170, loss: 0.03915967047214508
/usr/local/lib/python3.7/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=0.7305699481865284, f1=0.731829573934837,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07929570227861404
step: 10, loss: 0.004424956627190113
step: 20, loss: 0.029191771522164345
step: 30, loss: 0.0009593503782525659
step: 40, loss: 0.00027129179215990007
step: 50, loss: 0.04683028906583786

```

```

step: 60, loss: 0.0002323035878362134
step: 70, loss: 0.00047028460539877415
step: 80, loss: 0.0005807930720038712
step: 90, loss: 0.0014759802725166082
step: 100, loss: 0.0003390490310266614
step: 110, loss: 0.0032763127237558365
step: 120, loss: 0.0006977160228416324
step: 130, loss: 0.07081225514411926
step: 140, loss: 0.0294790156185627
step: 150, loss: 0.019237587228417397
step: 160, loss: 0.0013541651424020529
step: 170, loss: 0.00020318325550761074
/usr/local/lib/python3.7/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=0.7008086253369273, f1=0.7195767195767196,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016156090423464775
step: 10, loss: 0.04547657445073128
step: 20, loss: 0.0009994310094043612
step: 30, loss: 0.00018911277584265918
step: 40, loss: 0.00025773816742002964
step: 50, loss: 0.019446110352873802
step: 60, loss: 0.001477473066188395
step: 70, loss: 0.00012025044270558283
step: 80, loss: 0.0004950332804583013
step: 90, loss: 0.007754865568131208
step: 100, loss: 0.1434900015592575
step: 110, loss: 0.002059886697679758
step: 120, loss: 0.0008541603456251323
step: 130, loss: 0.00016806474013719708
step: 140, loss: 0.000349832815118134
step: 150, loss: 0.0008959340048022568
step: 160, loss: 0.0010713626397773623
step: 170, loss: 0.0008968728361651301
/usr/local/lib/python3.7/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=0.7008086253369273, f1=0.7324675324675325,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032441545045003295
step: 10, loss: 0.0033650521654635668

```



```

step: 20, loss: 0.0005659344606101513
step: 30, loss: 0.00038237194530665874
step: 40, loss: 0.0008568759076297283
step: 50, loss: 0.0005166362971067429
step: 60, loss: 0.001419323612935841
step: 70, loss: 0.056952573359012604
step: 80, loss: 0.00020488942391239107
step: 90, loss: 0.0011408027494326234
step: 100, loss: 0.017691651359200478
step: 110, loss: 0.00538815651088953
step: 120, loss: 0.0008717197342775762
step: 130, loss: 0.001847284846007824
step: 140, loss: 0.00147532531991601
step: 150, loss: 0.005149048753082752
step: 160, loss: 0.00027857880922965705
step: 170, loss: 0.000899118953384459
/usr/local/lib/python3.7/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=0.6925207756232686, f1=0.7335092348284961,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014245304046198726
step: 10, loss: 0.00029713864205405116
step: 20, loss: 0.011634009890258312
step: 30, loss: 0.00020910931925754994
step: 40, loss: 0.0005186396301724017
step: 50, loss: 0.0006417762488126755
step: 60, loss: 0.0006212581065483391
step: 70, loss: 0.0006116931326687336
step: 80, loss: 0.00048647765652276576
step: 90, loss: 0.00026430084835737944
step: 100, loss: 0.001608199905604124
step: 110, loss: 0.0002170923980884254
step: 120, loss: 0.010729474015533924
step: 130, loss: 0.00014275313878897578
step: 140, loss: 0.0001812096597859636
step: 150, loss: 0.0018702119123190641
step: 160, loss: 0.0021281293593347073
step: 170, loss: 0.0005780864739790559
/usr/local/lib/python3.7/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=0.7016574585635358, f1=0.7229551451187335,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009703070973046124
step: 10, loss: 0.00013368572399485856
step: 20, loss: 0.00017179767019115388
step: 30, loss: 0.005606450606137514
step: 40, loss: 0.00026607626932673156
step: 50, loss: 0.0030709232669323683
step: 60, loss: 9.346521983388811e-05
step: 70, loss: 0.0006806238670833409
step: 80, loss: 0.0004324323672335595
step: 90, loss: 0.0001332537067355588
step: 100, loss: 0.0017854004399850965
step: 110, loss: 0.00020862228120677173
step: 120, loss: 0.0012613695580512285
step: 130, loss: 0.0004795871500391513
step: 140, loss: 0.00014973235374782234
step: 150, loss: 0.0004921674844808877
step: 160, loss: 8.152038208208978e-05
step: 170, loss: 0.00011517975508468226
/usr/local/lib/python3.7/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=0.7055555555555555, f1=0.738544474393531,
best_f1=0.7294685990338163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031533758738078177
step: 10, loss: 0.0004957924829795957
step: 20, loss: 0.001799883204512298
step: 30, loss: 0.00029347019153647125
step: 40, loss: 0.0007361680036410689
step: 50, loss: 8.380127837881446e-05
step: 60, loss: 0.0015384610742330551
step: 70, loss: 0.00012401278945617378
step: 80, loss: 0.0049828733317554
step: 90, loss: 0.0005504398723132908
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.009405652061104774
step: 110, loss: 0.00016874952416401356
step: 120, loss: 0.025405706837773323
step: 130, loss: 0.00023142203281167895
step: 140, loss: 0.0003673938917927444
step: 150, loss: 0.0002480291877873242
step: 160, loss: 0.00021004959126003087
step: 170, loss: 0.00023631479416508228
/usr/local/lib/python3.7/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=0.7058823529411764, f1=0.738544474393531,
best_f1=0.7294685990338163
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, 235.97it/s]
load_f1 = 0.7589498806682576
real_f1 = 0.7517730496453902
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 193.78it/s]

```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6192374229431152
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.597267210483551
step: 20, loss: 0.5394448637962341
step: 30, loss: 0.5171781182289124
step: 40, loss: 0.36119452118873596
step: 50, loss: 0.17636150121688843
step: 60, loss: 0.23745065927505493
step: 70, loss: 0.20133274793624878
step: 80, loss: 0.08308546990156174
step: 90, loss: 0.06850156933069229
step: 100, loss: 0.010220480151474476
step: 110, loss: 0.15546591579914093
step: 120, loss: 0.05979273095726967
step: 130, loss: 0.05053778365254402
step: 140, loss: 0.04779361933469772
step: 150, loss: 0.09860529750585556
step: 160, loss: 0.011975212022662163
step: 170, loss: 0.21630778908729553
step: 180, loss: 0.2070765644311905
step: 190, loss: 0.1239706352353096
step: 200, loss: 0.10929886251688004
step: 210, loss: 0.02621220052242279
step: 220, loss: 0.1796300709247589
step: 230, loss: 0.09730477631092072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9401330376940134, f1=0.9361702127659575,
best_f1=0.9361702127659575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021602654829621315
step: 10, loss: 0.02358368970453739
step: 20, loss: 0.31484007835388184
step: 30, loss: 0.17102093994617462
step: 40, loss: 0.11015377938747406
step: 50, loss: 0.00829608179628849
step: 60, loss: 0.04810884967446327
step: 70, loss: 0.12798620760440826
step: 80, loss: 0.09878485649824142
step: 90, loss: 0.12859505414962769
step: 100, loss: 0.12216494232416153
step: 110, loss: 0.2065652757883072
step: 120, loss: 0.1433434933423996
step: 130, loss: 0.1100386530160904
step: 140, loss: 0.014153712429106236
step: 150, loss: 0.034401535987854004
step: 160, loss: 0.1118466705083847
step: 170, loss: 0.02172342874109745
step: 180, loss: 0.055773865431547165
step: 190, loss: 0.010463513433933258
step: 200, loss: 0.02588501013815403
step: 210, loss: 0.0026519515085965395
step: 220, loss: 0.032827313989400864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9419642857142857, f1=0.9228998849252014,
best_f1=0.9228998849252014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12792013585567474
step: 10, loss: 0.052666403353214264
step: 20, loss: 0.07030916213989258

```

```

step: 30, loss: 0.026775751262903214
step: 40, loss: 0.0072591411881148815
step: 50, loss: 0.019829895347356796
step: 60, loss: 0.006519021466374397
step: 70, loss: 0.18595822155475616
step: 80, loss: 0.0033091462682932615
step: 90, loss: 0.14166834950447083
step: 100, loss: 0.042549729347229004
step: 110, loss: 0.008393160067498684
step: 120, loss: 0.006145324092358351
step: 130, loss: 0.01033099927008152
step: 140, loss: 0.003221845254302025
step: 150, loss: 0.0263877771794796
step: 160, loss: 0.010580495931208134
step: 170, loss: 0.020861594006419182
step: 180, loss: 0.021061507984995842
step: 190, loss: 0.05702004209160805
step: 200, loss: 0.028562244027853012
step: 210, loss: 0.0041435775347054005
step: 220, loss: 0.051696110516786575
step: 230, loss: 0.009380968287587166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9500554938956715, f1=0.9414414414414415,
best_f1=0.9414414414414415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04623585194349289
step: 10, loss: 0.0016593532636761665
step: 20, loss: 0.015224194154143333
step: 30, loss: 0.005614478141069412
step: 40, loss: 0.004514869302511215
step: 50, loss: 0.002851094352081418
step: 60, loss: 0.004062525928020477
step: 70, loss: 0.0013076207833364606
step: 80, loss: 0.005530132912099361
step: 90, loss: 0.06208471208810806
step: 100, loss: 0.03966662660241127
step: 110, loss: 0.05122245103120804
step: 120, loss: 0.05910513922572136
step: 130, loss: 0.04087114706635475
step: 140, loss: 0.012831274420022964

```

```

step: 150, loss: 0.20316541194915771
step: 160, loss: 0.0080532506108284
step: 170, loss: 0.028359999880194664
step: 180, loss: 0.046071987599134445
step: 190, loss: 0.0010371883399784565
step: 200, loss: 0.0020332704298198223
step: 210, loss: 0.006648211739957333
step: 220, loss: 0.000638107187114656
step: 230, loss: 0.011221016757190228
/usr/local/lib/python3.7/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=0.9448275862068966, f1=0.9433526011560694,
best_f1=0.9414414414414415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010265017626807094
step: 10, loss: 0.00796507578343153
step: 20, loss: 0.005000915378332138
step: 30, loss: 0.0010109922150149941
step: 40, loss: 0.0008959921542555094
step: 50, loss: 0.0014531795168295503
step: 60, loss: 0.11828725785017014
step: 70, loss: 0.012392068281769753
step: 80, loss: 0.0006763305864296854
step: 90, loss: 0.048538390547037125
step: 100, loss: 0.0003404008166398853
step: 110, loss: 0.0003088300582021475
step: 120, loss: 0.01734551601111889
step: 130, loss: 0.08279100805521011
step: 140, loss: 0.009595402516424656
step: 150, loss: 0.0056921085342764854
step: 160, loss: 0.00500135775655508
step: 170, loss: 0.07389097660779953
step: 180, loss: 0.019160915166139603
step: 190, loss: 0.11215809732675552
step: 200, loss: 0.002160137752071023
step: 210, loss: 0.002600285457447171
step: 220, loss: 0.11974424868822098
step: 230, loss: 0.0010942231165245175
/usr/local/lib/python3.7/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=0.9484304932735426, f1=0.9446327683615819,
best_f1=0.9414414414414415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09845485538244247
step: 10, loss: 0.006768021732568741
step: 20, loss: 0.017974959686398506
step: 30, loss: 0.000449918006779626
step: 40, loss: 0.00038444335223175585
step: 50, loss: 0.006430627778172493
step: 60, loss: 0.0013551837764680386
step: 70, loss: 0.0019934263546019793
step: 80, loss: 0.002811969257891178
step: 90, loss: 0.0013906116364523768
step: 100, loss: 0.018982086330652237
step: 110, loss: 0.0010881503112614155
step: 120, loss: 0.018388858065009117
step: 130, loss: 0.0009881739970296621
step: 140, loss: 0.007716981694102287
step: 150, loss: 0.0006362285930663347
step: 160, loss: 0.0016765452455729246
step: 170, loss: 0.0003799708210863173
step: 180, loss: 0.006041708402335644
step: 190, loss: 0.027992000803351402
step: 200, loss: 0.00026403716765344143
step: 210, loss: 0.0007387488731183112
step: 220, loss: 0.0002579468709882349
step: 230, loss: 0.10463567823171616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9579646017699115, f1=0.9395973154362416,
best_f1=0.9395973154362416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06293725967407227
step: 10, loss: 0.0003571865672711283
step: 20, loss: 0.0007503183442167938
step: 30, loss: 0.0001985951530514285
step: 40, loss: 0.0008658805745653808
step: 50, loss: 0.00020192931697238237
step: 60, loss: 0.00041477399645373225
step: 70, loss: 0.009640965610742569
step: 80, loss: 0.05036235228180885
step: 90, loss: 0.00014702729822602123
step: 100, loss: 0.06660943478345871

```

```

step: 110, loss: 0.0003684107505250722
step: 120, loss: 0.0001698801788734272
step: 130, loss: 0.0063896626234054565
step: 140, loss: 0.0029183239676058292
step: 150, loss: 0.0006013571401126683
step: 160, loss: 0.07555725425481796
step: 170, loss: 0.00611201161518693
step: 180, loss: 0.01711127534508705
step: 190, loss: 0.00039019997348077595
step: 200, loss: 0.13991887867450714
step: 210, loss: 0.0005685937940143049
step: 220, loss: 0.0009862721199169755
step: 230, loss: 0.000745572498999536
/usr/local/lib/python3.7/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=0.9571106094808126, f1=0.9440000000000001,
best_f1=0.9395973154362416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016857525333762169
step: 10, loss: 0.0015629589324817061
step: 20, loss: 0.0004104462277609855
step: 30, loss: 0.000632257666438818
step: 40, loss: 0.017588673159480095
step: 50, loss: 0.001093595870770514
step: 60, loss: 0.0025640076491981745
step: 70, loss: 0.00033713155426084995
step: 80, loss: 0.0008136958349496126
step: 90, loss: 0.00015819842519704252
step: 100, loss: 0.0030260435305535793
step: 110, loss: 0.0007954707252793014
step: 120, loss: 0.05621642246842384
step: 130, loss: 0.03353030979633331
step: 140, loss: 0.0015323514817282557
step: 150, loss: 0.0006205764366313815
step: 160, loss: 0.0005882765399292111
step: 170, loss: 0.0014253866393119097
step: 180, loss: 0.0007258960977196693
step: 190, loss: 0.0012150516267865896
step: 200, loss: 0.009265446104109287
step: 210, loss: 0.06644199043512344
step: 220, loss: 0.0031312773935496807
step: 230, loss: 0.00042017639498226345
/usr/local/lib/python3.7/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=0.9425028184892897, f1=0.9419795221843003,
best_f1=0.9395973154362416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004044073633849621
step: 10, loss: 0.003407808020710945
step: 20, loss: 0.0011596346739679575
step: 30, loss: 0.0002273138816235587
step: 40, loss: 0.020944032818078995
step: 50, loss: 0.0001501339575042948
step: 60, loss: 0.000441274227341637
step: 70, loss: 0.0002123824378941208
step: 80, loss: 0.000956738309469074
step: 90, loss: 0.0005079020047560334
step: 100, loss: 0.0009632721194066107
step: 110, loss: 0.0001331191451754421
step: 120, loss: 0.000130232801893726
step: 130, loss: 0.0001529761211713776
step: 140, loss: 0.00011742604692699388
step: 150, loss: 0.007789317984133959
step: 160, loss: 0.0002031233161687851
step: 170, loss: 0.0008674691780470312
step: 180, loss: 0.0022025718353688717
step: 190, loss: 0.000742247502785176
step: 200, loss: 0.0040336549282073975
step: 210, loss: 0.0002557203406468034
step: 220, loss: 0.0006782071432098746
step: 230, loss: 0.00940820574760437
/usr/local/lib/python3.7/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=0.9536423841059603, f1=0.9443207126948775,
best_f1=0.9395973154362416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031633165781386197
step: 10, loss: 0.001020333031192422
step: 20, loss: 0.00018852333596441895
step: 30, loss: 0.00028047620435245335
step: 40, loss: 0.00030775374034419656
step: 50, loss: 0.0005671365070156753
step: 60, loss: 0.0004174507921561599
step: 70, loss: 0.000150100706377998
step: 80, loss: 0.00016171076276805252
step: 90, loss: 0.0005754553130827844
step: 100, loss: 0.00022946472745388746

```

```

step: 110, loss: 0.0031277891248464584
step: 120, loss: 0.00012183981743874028
step: 130, loss: 0.0008482764242216945
step: 140, loss: 0.03877493739128113
step: 150, loss: 0.04291857033967972
step: 160, loss: 0.000147237500641495
step: 170, loss: 0.00013186593423597515
step: 180, loss: 0.005449497140944004
step: 190, loss: 0.000629477493930608
step: 200, loss: 8.566284668631852e-05
step: 210, loss: 6.840858986834064e-05
step: 220, loss: 0.0001925275573739782
step: 230, loss: 0.0005229470552876592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9609810479375697, f1=0.941834451901566,
best_f1=0.941834451901566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016454188153147697
step: 10, loss: 0.0001328248472418636
step: 20, loss: 0.001993667334318161
step: 30, loss: 0.18534667789936066
step: 40, loss: 0.0008548739715479314
step: 50, loss: 0.00017562233551871032
step: 60, loss: 0.0008386599947698414
step: 70, loss: 0.0001702490699244663
step: 80, loss: 0.0001006187085295096
step: 90, loss: 0.00010980281513184309
step: 100, loss: 0.00013599637895822525
step: 110, loss: 0.0079466188326478
step: 120, loss: 0.00010537495109019801
step: 130, loss: 5.740886990679428e-05
step: 140, loss: 0.0011717929737642407
step: 150, loss: 0.00018308026483282447
step: 160, loss: 7.64666183385998e-05
step: 170, loss: 0.0017782084178179502
step: 180, loss: 0.0067823706194758415
step: 190, loss: 0.00010371541429776698
step: 200, loss: 0.00016169172886293381
step: 210, loss: 6.0131511418148875e-05
step: 220, loss: 8.91808231244795e-05

```

```

step: 230, loss: 0.00010720072168624029
/usr/local/lib/python3.7/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=0.9552572706935123, f1=0.9403824521934758,
best_f1=0.941834451901566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013515989121515304
step: 10, loss: 9.038137795869261e-05
step: 20, loss: 0.00013235240476205945
step: 30, loss: 0.000688319792971015
step: 40, loss: 0.0001225862215505913
step: 50, loss: 0.00018648534023668617
step: 60, loss: 0.00016315936227329075
step: 70, loss: 7.980535883689299e-05
step: 80, loss: 8.90537048690021e-05
step: 90, loss: 7.47829326428473e-05
step: 100, loss: 6.209911953192204e-05
step: 110, loss: 0.0001417384046362713
step: 120, loss: 0.00010983838001266122
step: 130, loss: 0.00015774730127304792
step: 140, loss: 0.00018030735373031348
step: 150, loss: 8.095982047962025e-05
step: 160, loss: 9.463392052566633e-05
step: 170, loss: 9.231571311829612e-05
step: 180, loss: 0.00017856888007372618
step: 190, loss: 0.0002249311946798116
step: 200, loss: 8.754820009926334e-05
step: 210, loss: 0.0001087379569071345
step: 220, loss: 0.00011644836922641844
step: 230, loss: 0.0004467648104764521
/usr/local/lib/python3.7/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=0.9573991031390134, f1=0.9379932356257046,
best_f1=0.941834451901566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009480954613536596
step: 10, loss: 0.0021253034938126802
step: 20, loss: 9.833250805968419e-05
step: 30, loss: 0.00020142982248216867
step: 40, loss: 0.0004406693624332547
step: 50, loss: 0.029667483642697334
step: 60, loss: 0.00021327304420992732

```

```

step: 70, loss: 0.10080042481422424
step: 80, loss: 0.00032595047377981246
step: 90, loss: 0.00011058411473641172
step: 100, loss: 0.013699379749596119
step: 110, loss: 0.0033505093306303024
step: 120, loss: 0.00021078884310554713
step: 130, loss: 0.0002603689208626747
step: 140, loss: 0.00015644678205717355
step: 150, loss: 0.00012998068996239454
step: 160, loss: 0.00016892097482923418
step: 170, loss: 0.0002329289127374068
step: 180, loss: 8.249421807704493e-05
step: 190, loss: 0.00012384731962811202
step: 200, loss: 7.468482363037765e-05
step: 210, loss: 4.073302625329234e-05
step: 220, loss: 0.0005345841054804623
step: 230, loss: 0.00013182869588490576
/usr/local/lib/python3.7/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=0.9584736251402918, f1=0.9427609427609427,
best_f1=0.941834451901566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017614694661460817
step: 10, loss: 0.00040846315096132457
step: 20, loss: 7.17362345312722e-05
step: 30, loss: 0.0001387672673445195
step: 40, loss: 0.0016822899924591184
step: 50, loss: 0.00031255726935341954
step: 60, loss: 7.026587991276756e-05
step: 70, loss: 0.0001254765666089952
step: 80, loss: 8.171355875674635e-05
step: 90, loss: 0.0003793720679823309
step: 100, loss: 9.080159361474216e-05
step: 110, loss: 7.912238652352244e-05
step: 120, loss: 3.4110522392438725e-05
step: 130, loss: 6.832400686107576e-05
step: 140, loss: 9.906468767439947e-05
step: 150, loss: 4.845934017794207e-05
step: 160, loss: 0.00018191781418863684
step: 170, loss: 0.00012453498493414372
step: 180, loss: 7.150598685257137e-05
step: 190, loss: 7.455565355485305e-05
step: 200, loss: 4.706968684331514e-05
step: 210, loss: 0.00015109045489225537
step: 220, loss: 0.001071673003025353

```

```

step: 230, loss: 0.0002249642857350409
/usr/local/lib/python3.7/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=0.9584736251402918, f1=0.9452513966480447,
best_f1=0.941834451901566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.435822484083474e-05
step: 10, loss: 4.1332386899739504e-05
step: 20, loss: 6.760745600331575e-05
step: 30, loss: 7.880419434513897e-05
step: 40, loss: 7.937342161312699e-05
step: 50, loss: 0.01917242258787155
step: 60, loss: 0.00020022675744257867
step: 70, loss: 0.0019284480949863791
step: 80, loss: 0.00019017753947991878
step: 90, loss: 0.00026357383467257023
step: 100, loss: 7.579351949971169e-05
step: 110, loss: 5.742721623391844e-05
step: 120, loss: 0.00015806735609658062
step: 130, loss: 6.054988261894323e-05
step: 140, loss: 8.573295053793117e-05
step: 150, loss: 0.00023206767218653113
step: 160, loss: 3.844323873636313e-05
step: 170, loss: 6.217490590643138e-05
step: 180, loss: 0.0001379170862492174
step: 190, loss: 0.0002006852300837636
step: 200, loss: 0.0005769772105850279
step: 210, loss: 0.025541165843605995
step: 220, loss: 0.002728105057030916
step: 230, loss: 6.136508454801515e-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=0.9594594594594594, f1=0.9423728813559321,
best_f1=0.941834451901566
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True

```

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:15, 159.99it/s]
load_f1 = 0.9608062709966406
real_f1 = 0.9588431590656283
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 192.28it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

```

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

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt

```



```
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt
```

```
# Step 3: run the matcher
```

```
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm bert-base-uncased \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

UserWarning)

step: 0, loss: 0.6312988996505737

step: 10, loss: 0.6120545268058777

step: 20, loss: 0.5849272608757019

step: 30, loss: 0.38294240832328796

step: 40, loss: 0.37277650833129883

step: 50, loss: 0.17793136835098267

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

step: 60, loss: 0.17053000628948212

step: 70, loss: 0.2606331706047058

step: 80, loss: 0.19594521820545197

step: 90, loss: 0.3203250765800476

step: 100, loss: 0.22131787240505219

step: 110, loss: 0.08648928254842758

step: 120, loss: 0.11615309864282608

step: 130, loss: 0.11199084669351578

step: 140, loss: 0.1105596199631691

step: 150, loss: 0.06963010877370834

step: 160, loss: 0.08014663308858871

step: 170, loss: 0.2961997389793396

step: 180, loss: 0.10005684196949005

step: 190, loss: 0.025555722415447235

step: 200, loss: 0.18905536830425262

step: 210, loss: 0.07123304903507233

step: 220, loss: 0.2086721658706665

step: 230, loss: 0.23081597685813904

step: 240, loss: 0.20672421157360077

step: 250, loss: 0.09918347001075745

step: 260, loss: 0.060434482991695404

step: 270, loss: 0.0143789853900671

step: 280, loss: 0.10697359591722488

step: 290, loss: 0.16587094962596893

step: 300, loss: 0.056003063917160034

step: 310, loss: 0.19659112393856049

step: 320, loss: 0.08217141032218933

step: 330, loss: 0.07505617290735245

step: 340, loss: 0.18367835879325867

step: 350, loss: 0.04619000852108002

step: 360, loss: 0.10206036269664764

step: 370, loss: 0.1216365322470665

step: 380, loss: 0.05186821147799492

step: 390, loss: 0.07757783681154251

step: 400, loss: 0.36001312732696533

step: 410, loss: 0.08900579810142517

step: 420, loss: 0.20354539155960083

step: 430, loss: 0.151458278298378

step: 440, loss: 0.08696899563074112

step: 450, loss: 0.006797574460506439

```

step: 460, loss: 0.06720329821109772
step: 470, loss: 0.13491079211235046
step: 480, loss: 0.18444105982780457
step: 490, loss: 0.14394702017307281
step: 500, loss: 0.06621420383453369
step: 510, loss: 0.10721459239721298
step: 520, loss: 0.0883471742272377
step: 530, loss: 0.006052887067198753
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.913003663003663, f1=0.8923499770957398,
best_f1=0.8923499770957398
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11305545270442963
step: 10, loss: 0.15229228138923645
step: 20, loss: 0.07153095304965973
step: 30, loss: 0.08310869336128235
step: 40, loss: 0.05879473686218262
step: 50, loss: 0.146795392036438
step: 60, loss: 0.08028260618448257
step: 70, loss: 0.06063029170036316
step: 80, loss: 0.12396625429391861
step: 90, loss: 0.1027035266160965
step: 100, loss: 0.028267860412597656
step: 110, loss: 0.01986747235059738
step: 120, loss: 0.04510263726115227
step: 130, loss: 0.09240365773439407
step: 140, loss: 0.0231282077729702
step: 150, loss: 0.1773117333650589
step: 160, loss: 0.06647161394357681
step: 170, loss: 0.23635746538639069
step: 180, loss: 0.009736026637256145
step: 190, loss: 0.08305622637271881
step: 200, loss: 0.022811289876699448
step: 210, loss: 0.18900223076343536
step: 220, loss: 0.08740471303462982
step: 230, loss: 0.059232961386442184
step: 240, loss: 0.05392644554376602
step: 250, loss: 0.13166972994804382
step: 260, loss: 0.08939006924629211
step: 270, loss: 0.2770560085773468

```

```

step: 280, loss: 0.19335998594760895
step: 290, loss: 0.03394879773259163
step: 300, loss: 0.14351342618465424
step: 310, loss: 0.041352011263370514
step: 320, loss: 0.1562167853116989
step: 330, loss: 0.13899211585521698
step: 340, loss: 0.08546307682991028
step: 350, loss: 0.011348743923008442
step: 360, loss: 0.07288603484630585
step: 370, loss: 0.25682732462882996
step: 380, loss: 0.08678384125232697
step: 390, loss: 0.13939814269542694
step: 400, loss: 0.04330313578248024
step: 410, loss: 0.03218091279268265
step: 420, loss: 0.02307196520268917
step: 430, loss: 0.22337114810943604
step: 440, loss: 0.1395205408334732
step: 450, loss: 0.12499570101499557
step: 460, loss: 0.025294912979006767
step: 470, loss: 0.027754249051213264
step: 480, loss: 0.20959994196891785
step: 490, loss: 0.05026626959443092
step: 500, loss: 0.41876840591430664
step: 510, loss: 0.03487929701805115
step: 520, loss: 0.16151748597621918
step: 530, loss: 0.04522519186139107
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9181258873639375, f1=0.9005681818181818,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031016001477837563
step: 10, loss: 0.029693765565752983
step: 20, loss: 0.09549643099308014
step: 30, loss: 0.1563584953546524
step: 40, loss: 0.04945315793156624
step: 50, loss: 0.19116738438606262
step: 60, loss: 0.04421960934996605
step: 70, loss: 0.013662501238286495
step: 80, loss: 0.002956658136099577
step: 90, loss: 0.030283527448773384

```

```
step: 100, loss: 0.007975936867296696
step: 110, loss: 0.044260863214731216
step: 120, loss: 0.0016222380800172687
step: 130, loss: 0.0026013976894319057
step: 140, loss: 0.011346772313117981
step: 150, loss: 0.13989698886871338
step: 160, loss: 0.04320318624377251
step: 170, loss: 0.06611044704914093
step: 180, loss: 0.05332259088754654
step: 190, loss: 0.07449214160442352
step: 200, loss: 0.11808319389820099
step: 210, loss: 0.0815882533788681
step: 220, loss: 0.07044297456741333
step: 230, loss: 0.1668318212032318
step: 240, loss: 0.00894970539957285
step: 250, loss: 0.13657288253307343
step: 260, loss: 0.04042649641633034
step: 270, loss: 0.018261082470417023
step: 280, loss: 0.22486862540245056
step: 290, loss: 0.008312651887536049
step: 300, loss: 0.013183430768549442
step: 310, loss: 0.043125275522470474
step: 320, loss: 0.05662455037236214
step: 330, loss: 0.03462713956832886
step: 340, loss: 0.012729895301163197
step: 350, loss: 0.08194644749164581
step: 360, loss: 0.0067095039412379265
step: 370, loss: 0.034975919872522354
step: 380, loss: 0.00768465967848897
step: 390, loss: 0.011450052261352539
step: 400, loss: 0.04957668483257294
step: 410, loss: 0.009186419658362865
step: 420, loss: 0.24368806183338165
step: 430, loss: 0.029198719188570976
step: 440, loss: 0.024410221725702286
step: 450, loss: 0.03795856237411499
step: 460, loss: 0.020546993240714073
step: 470, loss: 0.05856271833181381
step: 480, loss: 0.015570420771837234
step: 490, loss: 0.01077294535934925
step: 500, loss: 0.12738068401813507
step: 510, loss: 0.009545812383294106
step: 520, loss: 0.07107733935117722
step: 530, loss: 0.0717330276966095
/usr/local/lib/python3.7/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=0.9138012246820537, f1=0.8915318744053282,
```

```
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022538531571626663
step: 10, loss: 0.037278398871421814
step: 20, loss: 0.001016333349980414
step: 30, loss: 0.01953766494989395
step: 40, loss: 0.055764514952898026
step: 50, loss: 0.006387483328580856
step: 60, loss: 0.008199410513043404
step: 70, loss: 0.004025967791676521
step: 80, loss: 0.002371171023696661
step: 90, loss: 0.1686495840549469
step: 100, loss: 0.0039262366481125355
step: 110, loss: 0.06476486474275589
step: 120, loss: 0.0014123163418844342
step: 130, loss: 0.0007426138618029654
step: 140, loss: 0.0008113610674627125
step: 150, loss: 0.026251990348100662
step: 160, loss: 0.085244320333004
step: 170, loss: 0.0848381295800209
step: 180, loss: 0.0015504079638049006
step: 190, loss: 0.020991051569581032
step: 200, loss: 0.003955488558858633
step: 210, loss: 0.06352987885475159
step: 220, loss: 0.003663974115625024
step: 230, loss: 0.0045556253753602505
step: 240, loss: 0.00847413670271635
step: 250, loss: 0.05843978002667427
step: 260, loss: 0.0025386016350239515
step: 270, loss: 0.052015453577041626
step: 280, loss: 0.1287452131509781
step: 290, loss: 0.010530262254178524
step: 300, loss: 0.0008414026233367622
step: 310, loss: 0.008850193582475185
step: 320, loss: 0.006421247962862253
step: 330, loss: 0.01515323854982853
step: 340, loss: 0.01951196976006031
step: 350, loss: 0.05244719237089157
step: 360, loss: 0.05728236585855484
step: 370, loss: 0.007413393352180719
step: 380, loss: 0.0033177677541971207
step: 390, loss: 0.0032982905395329
step: 400, loss: 0.05648469552397728
step: 410, loss: 0.0015152875566855073
step: 420, loss: 0.14740674197673798
step: 430, loss: 0.04684118926525116
```

```

step: 440, loss: 0.013112715445458889
step: 450, loss: 0.020559515804052353
step: 460, loss: 0.010557946749031544
step: 470, loss: 0.03422988951206207
step: 480, loss: 0.11202484369277954
step: 490, loss: 0.006273996084928513
step: 500, loss: 0.006728064268827438
step: 510, loss: 0.026427708566188812
step: 520, loss: 0.0380120724439621
step: 530, loss: 0.0544663704931736
/usr/local/lib/python3.7/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=0.9145259224661373, f1=0.9014615747289015,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06367892771959305
step: 10, loss: 0.09834224730730057
step: 20, loss: 0.00909350998699665
step: 30, loss: 0.005762360990047455
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.14653746783733368
step: 50, loss: 0.002226076787337661
step: 60, loss: 0.11363694816827774
step: 70, loss: 0.01799924299120903
step: 80, loss: 0.008626868948340416
step: 90, loss: 0.0694662407040596
step: 100, loss: 0.008425022475421429
step: 110, loss: 0.0011746177915483713
step: 120, loss: 0.0984334871172905
step: 130, loss: 0.0005582415033131838
step: 140, loss: 0.0021649289410561323
step: 150, loss: 0.004416324198246002
step: 160, loss: 0.0007262262515723705
step: 170, loss: 0.019370675086975098
step: 180, loss: 0.001357095199637115
step: 190, loss: 0.03648838773369789
step: 200, loss: 0.006638785358518362
step: 210, loss: 0.027766967192292213
step: 220, loss: 0.011332766152918339
step: 230, loss: 0.02200070023536682
step: 240, loss: 0.005779515020549297
step: 250, loss: 0.07460721582174301
step: 260, loss: 0.008417056873440742
step: 270, loss: 0.0050382767803967
step: 280, loss: 0.010271595790982246

```

```

step: 290, loss: 0.21948400139808655
step: 300, loss: 0.024017583578824997
step: 310, loss: 0.002083300380036235
step: 320, loss: 0.07664375007152557
step: 330, loss: 0.05871360003948212
step: 340, loss: 0.0029011585284024477
step: 350, loss: 0.0007540606311522424
step: 360, loss: 0.0067685917019844055
step: 370, loss: 0.02860182337462902
step: 380, loss: 0.006291455589234829
step: 390, loss: 0.00044679566053673625
step: 400, loss: 0.009436710737645626
step: 410, loss: 0.012626893818378448
step: 420, loss: 0.0015561410691589117
step: 430, loss: 0.012783823534846306
step: 440, loss: 0.24412430822849274
step: 450, loss: 0.05246780440211296
step: 460, loss: 0.0031237625516951084
step: 470, loss: 0.0018155827419832349
step: 480, loss: 0.003912054002285004
step: 490, loss: 0.0006713010952807963
step: 500, loss: 0.00630223099142313
step: 510, loss: 0.048610568046569824
step: 520, loss: 0.01101086474955082
step: 530, loss: 0.05384765565395355
/usr/local/lib/python3.7/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=0.8980952380952382, f1=0.8864503816793893,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031991302967071533
step: 10, loss: 0.0005165531183592975
step: 20, loss: 0.12752148509025574
step: 30, loss: 0.004158813040703535
step: 40, loss: 0.004099155310541391
step: 50, loss: 0.0013380462769418955
step: 60, loss: 0.139240562915802
step: 70, loss: 0.015984104946255684
step: 80, loss: 0.00990466121584177
step: 90, loss: 0.009274952113628387
step: 100, loss: 0.0039835041388869286
step: 110, loss: 0.03128312528133392
step: 120, loss: 0.011097203008830547
step: 130, loss: 0.001387601369060576
step: 140, loss: 0.015628674998879433

```



```

step: 150, loss: 0.0003766172449104488
step: 160, loss: 0.0020157883409410715
step: 170, loss: 0.06548310816287994
step: 180, loss: 0.0017874727491289377
step: 190, loss: 0.0014144309097900987
step: 200, loss: 0.027173127979040146
step: 210, loss: 0.0012947152135893703
step: 220, loss: 0.0015712257008999586
step: 230, loss: 0.0004603344714269042
step: 240, loss: 0.003270823508501053
step: 250, loss: 0.0003764788853004575
step: 260, loss: 0.00022685667499899864
step: 270, loss: 0.15691280364990234
step: 280, loss: 0.003065458731725812
step: 290, loss: 0.0005753978621214628
step: 300, loss: 0.0026093740016222
step: 310, loss: 0.003978602588176727
step: 320, loss: 0.0035488714929670095
step: 330, loss: 0.007832012139260769
step: 340, loss: 0.06885653734207153
step: 350, loss: 0.03804619610309601
step: 360, loss: 0.01733354665338993
step: 370, loss: 0.017749378457665443
step: 380, loss: 0.009627070277929306
step: 390, loss: 0.005876166746020317
step: 400, loss: 0.02174387127161026
step: 410, loss: 0.015447588637471199
step: 420, loss: 0.11577188968658447
step: 430, loss: 0.00439140060916543
step: 440, loss: 0.0006742941914126277
step: 450, loss: 0.004556441679596901
step: 460, loss: 0.0003957226581405848
step: 470, loss: 0.0005109530175104737
step: 480, loss: 0.015412931330502033
step: 490, loss: 0.00045011434121988714
step: 500, loss: 0.00936537142843008
step: 510, loss: 0.0007825492648407817
step: 520, loss: 0.006670070346444845
step: 530, loss: 0.020505623891949654
/usr/local/lib/python3.7/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=0.9037106622827618, f1=0.886783514921838,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014739183825440705

```

step: 10, loss: 0.18337851762771606
step: 20, loss: 0.0011373715242370963
step: 30, loss: 0.011227548122406006
step: 40, loss: 0.0019879986066371202
step: 50, loss: 0.09355413913726807
step: 60, loss: 0.0015540615422651172
step: 70, loss: 0.004838928114622831
step: 80, loss: 0.059558916836977005
step: 90, loss: 0.007386140059679747
step: 100, loss: 0.000541446846909821
step: 110, loss: 0.0031261369585990906
step: 120, loss: 0.0002743565128184855
step: 130, loss: 0.0008884160197339952
step: 140, loss: 0.01066612172871828
step: 150, loss: 0.00771903432905674
step: 160, loss: 0.0013948369305580854
step: 170, loss: 0.0007170294993557036
step: 180, loss: 0.00030318062636069953
step: 190, loss: 0.20494507253170013
step: 200, loss: 0.010875605046749115
step: 210, loss: 0.012767067179083824
step: 220, loss: 0.0006340679829008877
step: 230, loss: 0.0018426422029733658
step: 240, loss: 0.003306909929960966
step: 250, loss: 0.002016800455749035
step: 260, loss: 0.00021471831132657826
step: 270, loss: 0.00021052395459264517
step: 280, loss: 0.0011779608903452754
step: 290, loss: 0.002100408310070634
step: 300, loss: 0.0028668008744716644
step: 310, loss: 0.0001383631315547973
step: 320, loss: 0.00012335697829257697
step: 330, loss: 0.0006424038438126445
step: 340, loss: 0.007422857917845249
step: 350, loss: 0.0031126709654927254
step: 360, loss: 0.003907713573426008
step: 370, loss: 0.000960119825322181
step: 380, loss: 0.003780276281759143
step: 390, loss: 0.0002546731266193092
step: 400, loss: 0.0016679356340318918
step: 410, loss: 0.0029660894069820642
step: 420, loss: 0.00012117554433643818
step: 430, loss: 0.00023283690097741783
step: 440, loss: 0.03919306769967079
step: 450, loss: 0.00364487711340189
step: 460, loss: 0.000489572761580348
step: 470, loss: 0.006237827241420746
step: 480, loss: 0.04295346513390541

```

step: 490, loss: 0.0015192005084827542
step: 500, loss: 0.0003341458213981241
step: 510, loss: 0.004332198761403561
step: 520, loss: 0.02803516574203968
step: 530, loss: 0.021037602797150612
/usr/local/lib/python3.7/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=0.9140698772426818, f1=0.8898508898508899,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001074297004379332
step: 10, loss: 0.003644264303147793
step: 20, loss: 0.0006642626249231398
step: 30, loss: 0.000665127532556653
step: 40, loss: 0.07342851161956787
step: 50, loss: 0.01611611433327198
step: 60, loss: 0.0007113544270396233
step: 70, loss: 0.001449744449928403
step: 80, loss: 0.10949136316776276
step: 90, loss: 0.0002530768688302487
step: 100, loss: 0.0018889844650402665
step: 110, loss: 0.0006997347227297723
step: 120, loss: 0.20848067104816437
step: 130, loss: 0.0008963398286141455
step: 140, loss: 0.0016909635160118341
step: 150, loss: 0.0015421988209709525
step: 160, loss: 0.0767889991402626
step: 170, loss: 0.015623362734913826
step: 180, loss: 0.0029089662712067366
step: 190, loss: 0.0017719202442094684
step: 200, loss: 0.00603528693318367
step: 210, loss: 0.0003927029320038855
step: 220, loss: 0.004190640524029732
step: 230, loss: 0.021710962057113647
step: 240, loss: 0.002057307865470648
step: 250, loss: 0.000249202101258561
step: 260, loss: 0.03199220821261406
step: 270, loss: 0.013421478681266308
step: 280, loss: 0.05119415372610092
step: 290, loss: 0.0011754874140024185
step: 300, loss: 0.002590422285720706
step: 310, loss: 0.00010209763422608376
step: 320, loss: 0.046603139489889145
step: 330, loss: 0.00012330309255048633
step: 340, loss: 0.003121984424069524

```

```

step: 350, loss: 0.016036562621593475
step: 360, loss: 0.0002512555511202663
step: 370, loss: 0.0001726622140267864
step: 380, loss: 0.0004898954066447914
step: 390, loss: 0.001714494195766747
step: 400, loss: 0.00010664881119737402
step: 410, loss: 0.00012071152741555125
step: 420, loss: 0.0005666168872267008
step: 430, loss: 0.009544582106173038
step: 440, loss: 0.01039167307317257
step: 450, loss: 0.00012258286005817354
step: 460, loss: 0.0015024684835225344
step: 470, loss: 0.00018620496848598123
step: 480, loss: 0.00015108197112567723
step: 490, loss: 0.0018328914884477854
step: 500, loss: 0.012814669869840145
step: 510, loss: 0.0001599111274117604
step: 520, loss: 0.002212973078712821
step: 530, loss: 8.926507871365175e-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=0.9170593779453347, f1=0.8934348239771647,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015057733980938792
step: 10, loss: 0.0009879773715510964
step: 20, loss: 0.0009467718773521483
step: 30, loss: 0.001502510509453714
step: 40, loss: 0.0011274563148617744
step: 50, loss: 7.710942736594006e-05
step: 60, loss: 5.604400939773768e-05
step: 70, loss: 6.517820293083787e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.006985502317547798
step: 90, loss: 0.00011166165495524183
step: 100, loss: 0.004533527418971062
step: 110, loss: 9.266142296837643e-05
step: 120, loss: 0.000684081285726279
step: 130, loss: 0.003267135936766863
step: 140, loss: 0.0003030105144716799
step: 150, loss: 7.93580838944763e-05
step: 160, loss: 0.0008808259153738618
step: 170, loss: 0.0026136855594813824
step: 180, loss: 6.203730299603194e-05
step: 190, loss: 0.00021458846458699554

```

```

step: 200, loss: 0.00021462092990987003
step: 210, loss: 0.00037100317422300577
step: 220, loss: 0.00019036269804928452
step: 230, loss: 0.0001881328789750114
step: 240, loss: 8.949844777816907e-05
step: 250, loss: 0.0003469125076662749
step: 260, loss: 0.004943342413753271
step: 270, loss: 6.117023440310732e-05
step: 280, loss: 0.0003319739771541208
step: 290, loss: 0.000868295319378376
step: 300, loss: 7.347852806560695e-05
step: 310, loss: 0.0287235826253891
step: 320, loss: 9.314106137026101e-05
step: 330, loss: 0.0003014870162587613
step: 340, loss: 8.045372669585049e-05
step: 350, loss: 0.00022762508888263255
step: 360, loss: 9.242469241144136e-05
step: 370, loss: 9.044050239026546e-05
step: 380, loss: 0.0013955117901787162
step: 390, loss: 0.009962080977857113
step: 400, loss: 0.0003450678486842662
step: 410, loss: 0.05169619619846344
step: 420, loss: 0.00035982075496576726
step: 430, loss: 0.0006077808793634176
step: 440, loss: 0.0006114688003435731
step: 450, loss: 0.004486437421292067
step: 460, loss: 0.001243386883288622
step: 470, loss: 0.00027869342011399567
step: 480, loss: 8.423202234553173e-05
step: 490, loss: 0.01492652390152216
step: 500, loss: 0.0012763573322445154
step: 510, loss: 0.009404348209500313
step: 520, loss: 8.276075823232532e-05
step: 530, loss: 0.0009393560467287898
/usr/local/lib/python3.7/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=0.9136960600375235, f1=0.895032802249297,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.8765887186164036e-05
step: 10, loss: 6.342010601656511e-05
step: 20, loss: 4.026537499157712e-05
step: 30, loss: 5.4461979743791744e-05
step: 40, loss: 0.005910925101488829
step: 50, loss: 0.010535500943660736

```

step: 60, loss: 0.00012686551781371236
step: 70, loss: 3.5180568374926224e-05
step: 80, loss: 4.174064815742895e-05
step: 90, loss: 5.2855670219287276e-05
step: 100, loss: 3.46331435139291e-05
step: 110, loss: 5.3820254834135994e-05
step: 120, loss: 0.00018047119374386966
step: 130, loss: 8.325683302246034e-05
step: 140, loss: 0.0002665536303538829
step: 150, loss: 9.198034967994317e-05
step: 160, loss: 6.329462485155091e-05
step: 170, loss: 0.00014977114915382117
step: 180, loss: 0.00017346924869343638
step: 190, loss: 0.004861611407250166
step: 200, loss: 0.000794108840636909
step: 210, loss: 0.00022404214541893452
step: 220, loss: 0.03133158013224602
step: 230, loss: 9.961240721167997e-05
step: 240, loss: 0.0007355646812357008
step: 250, loss: 0.0010716193355619907
step: 260, loss: 0.0004482263175304979
step: 270, loss: 3.6957186239305884e-05
step: 280, loss: 5.268466338748112e-05
step: 290, loss: 0.0012435272801667452
step: 300, loss: 0.12056494504213333
step: 310, loss: 6.980368198128417e-05
step: 320, loss: 0.07403085380792618
step: 330, loss: 3.3742257073754445e-05
step: 340, loss: 0.0005802945815958083
step: 350, loss: 0.00046765367733314633
step: 360, loss: 0.0016569859581068158
step: 370, loss: 0.000140256219310686
step: 380, loss: 0.0039482032880187035
step: 390, loss: 0.004320201929658651
step: 400, loss: 0.0018637534230947495
step: 410, loss: 0.00036406697472557425
step: 420, loss: 0.0005903568235225976
step: 430, loss: 6.979315367061645e-05
step: 440, loss: 0.0026210895739495754
step: 450, loss: 0.00010965659748762846
step: 460, loss: 0.0006659994833171368
step: 470, loss: 0.00015615923621226102
step: 480, loss: 4.356934368843213e-05
step: 490, loss: 6.907583156134933e-05
step: 500, loss: 0.009920618496835232
step: 510, loss: 0.00014972513599786907
step: 520, loss: 0.0012542230542749166
step: 530, loss: 3.978062159148976e-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=0.912, f1=0.8976303317535544, best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026862790808081627
step: 10, loss: 0.0015570848481729627
step: 20, loss: 7.837265002308413e-05
step: 30, loss: 9.737965592648834e-05
step: 40, loss: 3.780944098252803e-05
step: 50, loss: 9.712979954201728e-05
step: 60, loss: 0.0036227023229002953
step: 70, loss: 0.00039325476973317564
step: 80, loss: 0.0006416525575332344
step: 90, loss: 7.528246351284906e-05
step: 100, loss: 0.0004066140973009169
step: 110, loss: 7.765692134853452e-05
step: 120, loss: 4.860326225752942e-05
step: 130, loss: 0.0010638892417773604
step: 140, loss: 8.149279892677441e-05
step: 150, loss: 0.00012130776303820312
step: 160, loss: 4.605590584105812e-05
step: 170, loss: 3.3913744118763134e-05
step: 180, loss: 7.996754720807076e-05
step: 190, loss: 0.0004626686277333647
step: 200, loss: 0.0002533129299990833
step: 210, loss: 0.00018377319793216884
step: 220, loss: 3.1328832847066224e-05
step: 230, loss: 3.0560113373212516e-05
step: 240, loss: 3.0479410270345397e-05
step: 250, loss: 0.00019334348326083273
step: 260, loss: 8.379311475437135e-05
step: 270, loss: 0.00041128668817691505
step: 280, loss: 0.00015555144636891782
step: 290, loss: 0.00020994678197894245
step: 300, loss: 3.6234603612683713e-05
step: 310, loss: 0.00036331542651169
step: 320, loss: 0.0010765183251351118
step: 330, loss: 7.789513620082289e-05
step: 340, loss: 6.583805225091055e-05
step: 350, loss: 5.406378841144033e-05
step: 360, loss: 0.0002609499788377434
step: 370, loss: 0.0006208647391758859
step: 380, loss: 0.00011278628517175093
step: 390, loss: 7.112978346413001e-05
step: 400, loss: 5.076901652500965e-05
```

```

step: 410, loss: 0.000766577897593379
step: 420, loss: 0.0004513306193985045
step: 430, loss: 0.00037765741581097245
step: 440, loss: 0.0003368455800227821
step: 450, loss: 0.00014412702876143157
step: 460, loss: 0.00019777686975430697
step: 470, loss: 0.00015105344937182963
step: 480, loss: 0.00039235380245372653
step: 490, loss: 0.0001515205076429993
step: 500, loss: 0.0009484555921517313
step: 510, loss: 0.0005622321623377502
step: 520, loss: 0.00045590163790620863
step: 530, loss: 0.0006657494232058525
/usr/local/lib/python3.7/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=0.9092609915809168, f1=0.8987816307403936,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004645221401005983
step: 10, loss: 4.046285175718367e-05
step: 20, loss: 0.0004898622282780707
step: 30, loss: 5.073330248706043e-05
step: 40, loss: 4.5133947423892096e-05
step: 50, loss: 0.0006465790793299675
step: 60, loss: 4.187813465250656e-05
step: 70, loss: 0.00017257590661756694
step: 80, loss: 0.000264236907241866
step: 90, loss: 9.726163989398628e-05
step: 100, loss: 0.0004024810914415866
step: 110, loss: 0.0005763708031736314
step: 120, loss: 5.8288020227337256e-05
step: 130, loss: 0.0013672279892489314
step: 140, loss: 0.00019905628869310021
step: 150, loss: 4.9954171117860824e-05
step: 160, loss: 5.027649604016915e-05
step: 170, loss: 9.53081835177727e-05
step: 180, loss: 0.00033635986619628966
step: 190, loss: 0.0006336223450489342
step: 200, loss: 4.5923577999928966e-05
step: 210, loss: 4.805703429155983e-05
step: 220, loss: 0.0002836530329659581
step: 230, loss: 6.614446465391666e-05
step: 240, loss: 0.002964322455227375
step: 250, loss: 5.0782415200956166e-05
step: 260, loss: 5.3863335779169574e-05

```



```

step: 270, loss: 0.0001489875139668584
step: 280, loss: 3.0129400329315104e-05
step: 290, loss: 0.0001950897421920672
step: 300, loss: 0.00013402807235252112
step: 310, loss: 4.325179907027632e-05
step: 320, loss: 4.44449542555958e-05
step: 330, loss: 3.386866228538565e-05
step: 340, loss: 5.357724876375869e-05
step: 350, loss: 0.0012606356758624315
step: 360, loss: 0.0001356404391117394
step: 370, loss: 0.2376595288515091
step: 380, loss: 0.00038509684964083135
step: 390, loss: 4.961000377079472e-05
step: 400, loss: 0.0004480724746827036
step: 410, loss: 0.1288650929927826
step: 420, loss: 0.0034298799000680447
step: 430, loss: 0.00014916181680746377
step: 440, loss: 0.09266512095928192
step: 450, loss: 0.0107141537591815
step: 460, loss: 4.818346133106388e-05
step: 470, loss: 4.2176106944680214e-05
step: 480, loss: 0.0010080524953082204
step: 490, loss: 0.0032599212136119604
step: 500, loss: 7.87758908700198e-05
step: 510, loss: 0.0002624249318614602
step: 520, loss: 9.133601997746155e-05
step: 530, loss: 0.0005122878937982023
/usr/local/lib/python3.7/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=0.915764705882353, f1=0.8973995271867612,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015205575618892908
step: 10, loss: 0.0010598558001220226
step: 20, loss: 0.0002796074841171503
step: 30, loss: 0.00010317101987311617
step: 40, loss: 4.241799615556374e-05
step: 50, loss: 0.004288441967219114
step: 60, loss: 3.4650609450181946e-05
step: 70, loss: 3.689024015329778e-05
step: 80, loss: 0.00012051864905515686
step: 90, loss: 4.957336204824969e-05
step: 100, loss: 3.069572994718328e-05
step: 110, loss: 0.012793119065463543
step: 120, loss: 4.9776015657698736e-05

```

```

step: 130, loss: 3.51835296896752e-05
step: 140, loss: 6.874444079585373e-05
step: 150, loss: 0.00012230820721015334
step: 160, loss: 6.335152284009382e-05
step: 170, loss: 3.515445496304892e-05
step: 180, loss: 0.001122908084653318
step: 190, loss: 0.00010251544154016301
step: 200, loss: 4.970948430127464e-05
step: 210, loss: 3.575410300982185e-05
step: 220, loss: 5.4029278544476256e-05
step: 230, loss: 4.619134051608853e-05
step: 240, loss: 0.0005589399370364845
step: 250, loss: 4.783358963322826e-05
step: 260, loss: 6.233510066522285e-05
step: 270, loss: 3.9234673749888316e-05
step: 280, loss: 6.713434413541108e-05
step: 290, loss: 0.0011457018554210663
step: 300, loss: 7.177166844485328e-05
step: 310, loss: 0.00021797252702526748
step: 320, loss: 8.574975800001994e-05
step: 330, loss: 5.5514585255878046e-05
step: 340, loss: 5.6043329095700756e-05
step: 350, loss: 0.0001239521079696715
step: 360, loss: 5.3049800044391304e-05
step: 370, loss: 0.00018583284690976143
step: 380, loss: 2.77305025520036e-05
step: 390, loss: 0.0002981455181725323
step: 400, loss: 0.00014708892558701336
step: 410, loss: 0.00023758540919516236
step: 420, loss: 3.2297277357429266e-05
step: 430, loss: 0.0002620640443637967
step: 440, loss: 0.0003820652491413057
step: 450, loss: 2.7916750696022063e-05
step: 460, loss: 5.133591912453994e-05
step: 470, loss: 1.9721463104360737e-05
step: 480, loss: 2.1084866602905095e-05
step: 490, loss: 3.721046232385561e-05
step: 500, loss: 4.0240000089397654e-05
step: 510, loss: 3.612229193095118e-05
step: 520, loss: 0.0003546687075868249
step: 530, loss: 5.017447620048188e-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=0.9148737137511693, f1=0.9026798307475318,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.995739386184141e-05
step: 10, loss: 2.99317944154609e-05
step: 20, loss: 2.363651765335817e-05
step: 30, loss: 2.3960716134752147e-05
step: 40, loss: 2.5987150365835987e-05
step: 50, loss: 0.00035860075149685144
step: 60, loss: 1.9036020603380166e-05
step: 70, loss: 0.00048607756616547704
step: 80, loss: 9.005260653793812e-05
step: 90, loss: 4.4331431126920506e-05
step: 100, loss: 4.801843169843778e-05
step: 110, loss: 3.688264041556977e-05
step: 120, loss: 0.001516450080089271
step: 130, loss: 0.008664865046739578
step: 140, loss: 3.219622522010468e-05
step: 150, loss: 2.26606207434088e-05
step: 160, loss: 0.00014763603394385427
step: 170, loss: 2.5152703528874554e-05
step: 180, loss: 3.2990239560604095e-05
step: 190, loss: 2.9313341656234115e-05
step: 200, loss: 6.312198820523918e-05
step: 210, loss: 2.72871257038787e-05
step: 220, loss: 0.00018930000078398734
step: 230, loss: 9.603664511814713e-05
step: 240, loss: 3.4792399674188346e-05
step: 250, loss: 3.066578574362211e-05
step: 260, loss: 3.542809281498194e-05
step: 270, loss: 2.1844803995918483e-05
step: 280, loss: 2.089861845888663e-05
step: 290, loss: 8.032674668356776e-05
step: 300, loss: 0.00010561551607679576
step: 310, loss: 6.0356462199706584e-05
step: 320, loss: 0.00034044080530293286
step: 330, loss: 7.580497913295403e-05
step: 340, loss: 2.669488821993582e-05
step: 350, loss: 0.012409936636686325
step: 360, loss: 2.487320489308331e-05
step: 370, loss: 3.7753747164970264e-05
step: 380, loss: 2.8106742320233025e-05
step: 390, loss: 0.00949680432677269
step: 400, loss: 2.464611316099763e-05
step: 410, loss: 2.1610085241263732e-05
step: 420, loss: 3.728849333128892e-05
step: 430, loss: 5.1521798013709486e-05
step: 440, loss: 7.800190360285342e-05
step: 450, loss: 0.00010982140520354733
step: 460, loss: 2.3200760551844724e-05
```

```

step: 470, loss: 2.974084964080248e-05
step: 480, loss: 3.5727447539102286e-05
step: 490, loss: 3.0456463719019666e-05
step: 500, loss: 1.942341805261094e-05
step: 510, loss: 0.0001479663624195382
step: 520, loss: 3.785390799748711e-05
step: 530, loss: 2.3208205675473437e-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=0.9136385087305333, f1=0.9012753897024091,
best_f1=0.9005681818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2749962226953357e-05
step: 10, loss: 8.594375685788691e-05
step: 20, loss: 3.107521843048744e-05
step: 30, loss: 7.229580660350621e-05
step: 40, loss: 3.539257886586711e-05
step: 50, loss: 0.00034811344812624156
step: 60, loss: 1.9449485989753157e-05
step: 70, loss: 2.4444854716421105e-05
step: 80, loss: 2.6273844923707657e-05
step: 90, loss: 0.00017293334531132132
step: 100, loss: 0.00010076842590933666
step: 110, loss: 2.9603959774249233e-05
step: 120, loss: 3.578310861485079e-05
step: 130, loss: 0.005453836172819138
step: 140, loss: 2.0008274077554233e-05
step: 150, loss: 3.374975858605467e-05
step: 160, loss: 2.6344543584855273e-05
step: 170, loss: 2.222845796495676e-05
step: 180, loss: 1.676736428635195e-05
step: 190, loss: 2.537622094678227e-05
step: 200, loss: 4.5839806261938065e-05
step: 210, loss: 3.5625998862087727e-05
step: 220, loss: 2.5334811653010547e-05
step: 230, loss: 7.845590153010562e-05
step: 240, loss: 2.3100174075807445e-05
step: 250, loss: 7.115996413631365e-05
step: 260, loss: 2.2388676370610483e-05
step: 270, loss: 2.3133656213758513e-05
step: 280, loss: 2.2701527996105142e-05
step: 290, loss: 2.0693718397524208e-05
step: 300, loss: 3.365152588230558e-05
step: 310, loss: 3.8260684959823266e-05
step: 320, loss: 7.941704825498164e-05

```

```

step: 330, loss: 5.558469638344832e-05
step: 340, loss: 2.4269786081276834e-05
step: 350, loss: 1.9289225747343153e-05
step: 360, loss: 3.98768279410433e-05
step: 370, loss: 2.2295422240858898e-05
step: 380, loss: 2.235141619166825e-05
step: 390, loss: 4.6660996304126456e-05
step: 400, loss: 2.2690406694891863e-05
step: 410, loss: 0.0036327382549643517
step: 420, loss: 2.0924608179484494e-05
step: 430, loss: 8.229284139815718e-05
step: 440, loss: 9.134544234257191e-05
step: 450, loss: 3.977260712417774e-05
step: 460, loss: 2.3066608264343813e-05
step: 470, loss: 4.8712314310250804e-05
step: 480, loss: 1.987043651752174e-05
step: 490, loss: 2.1788908270536922e-05
step: 500, loss: 7.130094309104607e-05
step: 510, loss: 3.084295894950628e-05
step: 520, loss: 2.3621614673174918e-05
step: 530, loss: 2.497751484042965e-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=0.9140733859730609, f1=0.9010782934833569,
best_f1=0.9005681818181818
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:30, 188.59it/s]
load_f1 = 0.9195075757575758
real_f1 = 0.9152863227638429
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 190.05it/s]

```

3 DK

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                               | 827.9 MB 1.3 MB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: plac<1.2.0,>=0.9.6 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.1.3)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: 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: 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: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)

Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)

Requirement already satisfied: 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: 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: 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: 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: 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: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)

Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)

Requirement already satisfied: 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: 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: 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=76eb5becc83d5a086bb88d38716bd9135cedbf72058d0cf6c3559295df529a43
  Stored in directory: /tmp/pip-ephem-wheel-cache-h2l3tupy/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

```

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

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

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

```

Downloading: 100% 433/433 [00:00<00:00, 457kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 3.77MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 68.8MB/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: 0.5580651760101318

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

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

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

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=0.3684210526315789, f1=0.33766233766233766,
best_f1=0.33766233766233766

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.487705796957016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5555555555555556, f1=0.48, best_f1=0.48
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48420560359954834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.5, f1=0.44067796610169485, best_f1=0.48
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3797960877418518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6222222222222222, f1=0.4905660377358491,
best_f1=0.4905660377358491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25421980023384094
/usr/local/lib/python3.7/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=0.6060606060606061, f1=0.5454545454545454,
best_f1=0.4905660377358491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34076476097106934
/usr/local/lib/python3.7/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=0.5384615384615384, f1=0.6206896551724138,
best_f1=0.4905660377358491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3629540503025055
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6486486486486486, f1=0.55, best_f1=0.55
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14020290970802307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.75, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008622473105788231
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.5384615384615384, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12401348352432251
/usr/local/lib/python3.7/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=0.6956521739130435, f1=0.5384615384615384, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04934275895357132
/usr/local/lib/python3.7/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=0.7200000000000001, f1=0.6000000000000001, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04014194756746292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7741935483870968, f1=0.6060606060606061,
best_f1=0.6060606060606061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03465431183576584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7878787878787878, f1=0.6060606060606061,
best_f1=0.6060606060606061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047165825963020325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8125000000000001, f1=0.6060606060606061,
best_f1=0.6060606060606061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07806573808193207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8125000000000001, f1=0.6060606060606061,
best_f1=0.6060606060606061
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, 133923.39it/s]
/usr/local/lib/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 = 0.7058823529411764
real_f1 = 0.6857142857142857
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 257.01it/s]

```

3.1.2 DBLP-ACM - Running the matcher

```

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

```

```

# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-ACM \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm bert-base-uncased \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6152023673057556
step: 10, loss: 0.642973005771637
step: 20, loss: 0.47908568382263184
step: 30, loss: 0.33732637763023376
step: 40, loss: 0.22306329011917114
step: 50, loss: 0.16794683039188385
step: 60, loss: 0.12328655272722244
step: 70, loss: 0.03905951604247093
step: 80, loss: 0.06941882520914078
step: 90, loss: 0.04841301962733269
step: 100, loss: 0.004194248467683792
step: 110, loss: 0.2970150411128998
step: 120, loss: 0.04597296193242073
step: 130, loss: 0.029168453067541122
step: 140, loss: 0.00946805253624916
step: 150, loss: 0.07201418280601501
step: 160, loss: 0.020635556429624557
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.20727062225341797
step: 180, loss: 0.16266989707946777
step: 190, loss: 0.04804529994726181
step: 200, loss: 0.023387182503938675
step: 210, loss: 0.05062030628323555
step: 220, loss: 0.02660377323627472
step: 230, loss: 0.05675648897886276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9719416386083053, f1=0.9706546275395034,
best_f1=0.9706546275395034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1134919673204422
step: 10, loss: 0.006605489645153284
step: 20, loss: 0.10926666110754013
step: 30, loss: 0.18808215856552124
step: 40, loss: 0.07511946558952332
step: 50, loss: 0.01607067696750164
step: 60, loss: 0.011250887997448444
step: 70, loss: 0.18302884697914124
step: 80, loss: 0.007688598241657019
step: 90, loss: 0.005607320461422205
step: 100, loss: 0.006958906073123217
step: 110, loss: 0.15910880267620087
step: 120, loss: 0.0020396483596414328
step: 130, loss: 0.010465122759342194
step: 140, loss: 0.026741860434412956
step: 150, loss: 0.09666755050420761
step: 160, loss: 0.035107821226119995
step: 170, loss: 0.0034125763922929764
step: 180, loss: 0.0050445920787751675
step: 190, loss: 0.07382607460021973
step: 200, loss: 0.008446523919701576
step: 210, loss: 0.0010954281315207481
step: 220, loss: 0.05454663187265396
step: 230, loss: 0.02339094877243042
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9766407119021134, f1=0.9688195991091313,
best_f1=0.9688195991091313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012628396973013878
step: 10, loss: 0.012472317554056644
step: 20, loss: 0.005277307238429785
step: 30, loss: 0.0015672520967200398
step: 40, loss: 0.08124203234910965

```



```
step: 50, loss: 0.09487494826316833
step: 60, loss: 0.000975185539573431
step: 70, loss: 0.05059468746185303
step: 80, loss: 0.0003397518885321915
step: 90, loss: 0.15499071776866913
step: 100, loss: 0.000493288564030081
step: 110, loss: 0.0021832669153809547
step: 120, loss: 0.03289487585425377
step: 130, loss: 0.003403279697522521
step: 140, loss: 0.04337836429476738
step: 150, loss: 0.06494028121232986
step: 160, loss: 0.008172111585736275
step: 170, loss: 0.0033379453234374523
step: 180, loss: 0.10597190260887146
step: 190, loss: 0.03616920858621597
step: 200, loss: 0.024507025256752968
step: 210, loss: 0.0005018914234824479
step: 220, loss: 0.0010960126528516412
step: 230, loss: 0.0010388686787337065
/usr/local/lib/python3.7/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=0.9659090909090909, f1=0.9655963302752294,
best_f1=0.9688195991091313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030850248876959085
step: 10, loss: 0.0003233612223993987
step: 20, loss: 0.0002934943186119199
step: 30, loss: 0.0001795912248780951
step: 40, loss: 0.0012395278317853808
step: 50, loss: 0.0014830139698460698
step: 60, loss: 0.008968710899353027
step: 70, loss: 0.0017852711025625467
step: 80, loss: 0.007622533943504095
step: 90, loss: 0.024081028997898102
step: 100, loss: 0.0043403636664152145
step: 110, loss: 0.0014084617141634226
step: 120, loss: 0.014088444411754608
step: 130, loss: 0.0011082038981840014
step: 140, loss: 0.0009416933753527701
step: 150, loss: 0.1553025245666504
step: 160, loss: 0.03940892592072487
step: 170, loss: 0.028937172144651413
step: 180, loss: 0.0009928479557856917
step: 190, loss: 0.04063193500041962
step: 200, loss: 0.0013092408189550042
```

```

step: 210, loss: 0.12553507089614868
step: 220, loss: 0.003965489566326141
step: 230, loss: 0.048259977251291275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9796380090497738, f1=0.9808773903262092,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002573311096057296
step: 10, loss: 0.0029836164321750402
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.045654501765966415
step: 30, loss: 0.00029814179288223386
step: 40, loss: 0.0003422588633839041
step: 50, loss: 0.0002155730762751773
step: 60, loss: 0.0002966538304463029
step: 70, loss: 0.00019582251843530685
step: 80, loss: 0.0004470002604648471
step: 90, loss: 0.000784163421485573
step: 100, loss: 0.00029462159727700055
step: 110, loss: 0.0002236179861938581
step: 120, loss: 0.00020779127953574061
step: 130, loss: 0.0004684515006374568
step: 140, loss: 0.08108212798833847
step: 150, loss: 0.1367213875055313
step: 160, loss: 0.005343434400856495
step: 170, loss: 0.012113839387893677
step: 180, loss: 0.019560175016522408
step: 190, loss: 0.027233976870775223
step: 200, loss: 0.0014534093206748366
step: 210, loss: 0.0014427014393731952
step: 220, loss: 0.004398318938910961
step: 230, loss: 0.00267969723790884
/usr/local/lib/python3.7/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=0.9729119638826186, f1=0.972972972972973,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.006143142003566027
step: 10, loss: 0.0022012023255228996
step: 20, loss: 0.039387594908475876
step: 30, loss: 0.00027870663325302303
step: 40, loss: 0.0009535423596389592
step: 50, loss: 0.00024771972675807774
step: 60, loss: 0.00027513678651303053
step: 70, loss: 0.0005399013170972466
step: 80, loss: 0.05333775654435158
step: 90, loss: 0.04187118634581566
step: 100, loss: 0.04811769723892212
step: 110, loss: 0.0026248206850141287
step: 120, loss: 0.0224298182874918
step: 130, loss: 0.0017983062425628304
step: 140, loss: 0.007895401678979397
step: 150, loss: 0.01155738066881895
step: 160, loss: 0.10109933465719223
step: 170, loss: 0.00046728577581234276
step: 180, loss: 0.03788599371910095
step: 190, loss: 0.011873210780322552
step: 200, loss: 0.0004582286346703768
step: 210, loss: 0.09045626223087311
step: 220, loss: 0.0004492232983466238
step: 230, loss: 0.10660668462514877
/usr/local/lib/python3.7/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=0.9764309764309763, f1=0.9743589743589743,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003284905571490526
step: 10, loss: 0.00042978659621439874
step: 20, loss: 0.0012761215912178159
step: 30, loss: 0.00771896680817008
step: 40, loss: 0.0005047654267400503
step: 50, loss: 0.00012247431732248515
step: 60, loss: 0.00011421951785450801
step: 70, loss: 0.012532681226730347
step: 80, loss: 0.00021310480951797217
step: 90, loss: 0.0003992079582531005
step: 100, loss: 0.001272610854357481
step: 110, loss: 0.001944702584296465
step: 120, loss: 0.0002757494803518057
step: 130, loss: 0.0001529179426142946
step: 140, loss: 0.0005087650497443974
step: 150, loss: 0.0017699608579277992

```

```

step: 160, loss: 0.06661542505025864
step: 170, loss: 0.002014779020100832
step: 180, loss: 0.00021507182100322098
step: 190, loss: 0.0003153465804643929
step: 200, loss: 0.0018056684639304876
step: 210, loss: 6.079454760765657e-05
step: 220, loss: 0.00020462799875531346
step: 230, loss: 0.00013255832891445607
/usr/local/lib/python3.7/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=0.979591836734694, f1=0.9740698985343857,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.800145663670264e-05
step: 10, loss: 0.00013921638310421258
step: 20, loss: 0.0048362030647695065
step: 30, loss: 0.0002514634979888797
step: 40, loss: 0.0002439594391034916
step: 50, loss: 0.00031594137544743717
step: 60, loss: 0.00036289324634708464
step: 70, loss: 0.019082706421613693
step: 80, loss: 0.04907650873064995
step: 90, loss: 0.0001769417867762968
step: 100, loss: 0.00013431439583655447
step: 110, loss: 0.0001322216121479869
step: 120, loss: 0.0013449226971715689
step: 130, loss: 4.898485349258408e-05
step: 140, loss: 5.9835248976014555e-05
step: 150, loss: 0.0004924678360112011
step: 160, loss: 9.836194658419117e-05
step: 170, loss: 7.570208981633186e-05
step: 180, loss: 0.0005328080151230097
step: 190, loss: 0.00011477332009235397
step: 200, loss: 0.00014801652287133038
step: 210, loss: 8.838741632644087e-05
step: 220, loss: 9.100775787374005e-05
step: 230, loss: 0.00016850886458996683
/usr/local/lib/python3.7/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=0.9775784753363228, f1=0.9730337078651685,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.07272788137197495
step: 10, loss: 0.0030842386186122894
step: 20, loss: 0.0009928279323503375
step: 30, loss: 0.00017563486471772194
step: 40, loss: 0.004487267229706049
step: 50, loss: 6.255198240978643e-05
step: 60, loss: 0.0002481777628418058
step: 70, loss: 0.0007157461950555444
step: 80, loss: 0.006758770439773798
step: 90, loss: 0.000919120735488832
step: 100, loss: 0.03649481385946274
step: 110, loss: 0.00019266425806563348
step: 120, loss: 3.92221991205588e-05
step: 130, loss: 0.0002457774244248867
step: 140, loss: 0.00019773068197537214
step: 150, loss: 0.0008940441184677184
step: 160, loss: 0.00012201886420371011
step: 170, loss: 0.00016972990124486387
step: 180, loss: 0.007404760457575321
step: 190, loss: 0.0002954124065581709
step: 200, loss: 0.00018152117263525724
step: 210, loss: 0.00018214517331216484
step: 220, loss: 0.0005048344028182328
step: 230, loss: 0.00011026555148418993
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9763779527559054, f1=0.972972972972973,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006742983241565526
step: 10, loss: 7.472455035895109e-05
step: 20, loss: 0.00013119909272063524
step: 30, loss: 0.00012759093078784645
step: 40, loss: 5.3588988521369174e-05
step: 50, loss: 4.4444663217291236e-05
step: 60, loss: 8.30345816211775e-05
step: 70, loss: 0.00022432075638789684
step: 80, loss: 7.058295159367844e-05
step: 90, loss: 0.00030447158496826887
step: 100, loss: 0.0001922132942127064
step: 110, loss: 7.928773993626237e-05
step: 120, loss: 7.073193410178646e-05
step: 130, loss: 6.856227992102504e-05
step: 140, loss: 0.08114824444055557
step: 150, loss: 0.007622345816344023

```

```

step: 160, loss: 0.001475038006901741
step: 170, loss: 3.802268111030571e-05
step: 180, loss: 7.96044259914197e-05
step: 190, loss: 0.00154072733130306
step: 200, loss: 3.889799700118601e-05
step: 210, loss: 5.489652539836243e-05
step: 220, loss: 5.406217314884998e-05
step: 230, loss: 0.00021309168369043618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9752252252252253,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.1365022954996675e-05
step: 10, loss: 6.244950782274827e-05
step: 20, loss: 0.005470974836498499
step: 30, loss: 0.011929260566830635
step: 40, loss: 6.261100497795269e-05
step: 50, loss: 9.209438576363027e-05
step: 60, loss: 6.87385763740167e-05
step: 70, loss: 7.898407784523442e-05
step: 80, loss: 4.335646008257754e-05
step: 90, loss: 3.966520671383478e-05
step: 100, loss: 5.045926081947982e-05
step: 110, loss: 0.00039494462544098496
step: 120, loss: 0.00011087229358963668
step: 130, loss: 0.00013797437713947147
step: 140, loss: 0.0001681863359408453
step: 150, loss: 0.0017929155146703124
step: 160, loss: 0.00010209815081907436
step: 170, loss: 0.0004924061358906329
step: 180, loss: 9.260993829229847e-05
step: 190, loss: 0.00011227049981243908
step: 200, loss: 9.325963037554175e-05
step: 210, loss: 6.600052438443527e-05
step: 220, loss: 6.399302219506353e-05
step: 230, loss: 0.00011007746070390567
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9729119638826186,

```

```

best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.915057045058347e-05
step: 10, loss: 0.0002853100304491818
step: 20, loss: 6.685442349407822e-05
step: 30, loss: 7.472898141713813e-05
step: 40, loss: 0.0001390490069752559
step: 50, loss: 0.008390413597226143
step: 60, loss: 0.00017489930905867368
step: 70, loss: 6.552245758939534e-05
step: 80, loss: 7.874878065194935e-05
step: 90, loss: 4.634031211026013e-05
step: 100, loss: 0.00030970462830737233
step: 110, loss: 0.0001400786713929847
step: 120, loss: 5.4274467402137816e-05
step: 130, loss: 8.080504630925134e-05
step: 140, loss: 6.460761505877599e-05
step: 150, loss: 6.719423254253343e-05
step: 160, loss: 7.306145562324673e-05
step: 170, loss: 7.052408182062209e-05
step: 180, loss: 6.0975751694059e-05
step: 190, loss: 0.00012390484334900975
step: 200, loss: 4.1833081922959536e-05
step: 210, loss: 0.00011122236173832789
step: 220, loss: 4.217226160108112e-05
step: 230, loss: 6.214497261680663e-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=0.9796380090497738, f1=0.9739524348810873,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.983011113945395e-05
step: 10, loss: 0.0005079961847513914
step: 20, loss: 9.91438064374961e-05
step: 30, loss: 7.08340885466896e-05
step: 40, loss: 6.413453229470178e-05
step: 50, loss: 8.180810982594267e-05
step: 60, loss: 6.619618216063827e-05
step: 70, loss: 0.0001266994804609567
step: 80, loss: 7.106044358806685e-05
step: 90, loss: 0.0011031817412003875
step: 100, loss: 3.9755035686539486e-05
step: 110, loss: 0.1529414802789688

```

```

step: 120, loss: 0.0008784226956777275
step: 130, loss: 0.00017987797036767006
step: 140, loss: 6.304193811956793e-05
step: 150, loss: 5.453899211715907e-05
step: 160, loss: 0.002840946661308408
step: 170, loss: 0.00010445933730807155
step: 180, loss: 8.73687895364128e-05
step: 190, loss: 5.8500059822108597e-05
step: 200, loss: 0.00317385490052402
step: 210, loss: 7.661611743969843e-05
step: 220, loss: 0.0014230618253350258
step: 230, loss: 0.00010729907808126882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9796380090497738, f1=0.9728506787330317,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.587126386468299e-05
step: 10, loss: 0.009114964865148067
step: 20, loss: 8.548131154384464e-05
step: 30, loss: 0.0001548653090139851
step: 40, loss: 0.0010277745313942432
step: 50, loss: 0.00020910415332764387
step: 60, loss: 0.0001709893549559638
step: 70, loss: 0.0023769736289978027
step: 80, loss: 4.441106284502894e-05
step: 90, loss: 7.684188312850893e-05
step: 100, loss: 7.553654722869396e-05
step: 110, loss: 6.204845703905448e-05
step: 120, loss: 2.6586887543089688e-05
step: 130, loss: 4.586721115629189e-05
step: 140, loss: 5.5406246247002855e-05
step: 150, loss: 0.0011005926644429564
step: 160, loss: 0.0033007788006216288
step: 170, loss: 4.519377398537472e-05
step: 180, loss: 7.71130362409167e-05
step: 190, loss: 5.778158811153844e-05
step: 200, loss: 4.82863761135377e-05
step: 210, loss: 0.00011884041305165738
step: 220, loss: 4.070364229846746e-05
step: 230, loss: 0.000103725069493521
/usr/local/lib/python3.7/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=0.9751693002257337, f1=0.9683257918552037,

```



```

best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.6067081964574754e-05
step: 10, loss: 3.130986442556605e-05
step: 20, loss: 0.00045350557775236666
step: 30, loss: 0.00014837904018349946
step: 40, loss: 0.00020012620370835066
step: 50, loss: 0.04022315517067909
step: 60, loss: 0.0001912261504912749
step: 70, loss: 0.00010620113607728854
step: 80, loss: 5.0669183110585436e-05
step: 90, loss: 0.0001329927472397685
step: 100, loss: 5.336863250704482e-05
step: 110, loss: 0.0002086762833641842
step: 120, loss: 4.3930569518124685e-05
step: 130, loss: 8.294328290503472e-05
step: 140, loss: 3.7166060792515054e-05
step: 150, loss: 0.00020898182992823422
step: 160, loss: 3.67851389455609e-05
step: 170, loss: 3.2226416806224734e-05
step: 180, loss: 0.001550908898934722
step: 190, loss: 0.00037626922130584717
step: 200, loss: 0.00012531243555713445
step: 210, loss: 3.884606121573597e-05
step: 220, loss: 0.00045123809832148254
step: 230, loss: 4.257042019162327e-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=0.9761634506242906, f1=0.9704545454545453,
best_f1=0.9752252252252253
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:15, 158.73it/s]
load_f1 = 0.9797297297297298
real_f1 = 0.9751693002257337
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 183.03it/s]

```

3.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

step: 0, loss: 0.6203928589820862
step: 10, loss: 0.5956630706787109
step: 20, loss: 0.5956918001174927
step: 30, loss: 0.3407902717590332
step: 40, loss: 0.21666613221168518
step: 50, loss: 0.2774430215358734
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.1029977798461914
step: 70, loss: 0.19061754643917084
step: 80, loss: 0.04065834730863571
step: 90, loss: 0.4285522699356079
step: 100, loss: 0.3212401270866394
step: 110, loss: 0.05595703050494194
step: 120, loss: 0.17775672674179077
step: 130, loss: 0.19709612429141998
step: 140, loss: 0.12181587517261505
step: 150, loss: 0.1851647049188614
step: 160, loss: 0.0774395614862442
step: 170, loss: 0.16738298535346985
step: 180, loss: 0.06013816222548485
step: 190, loss: 0.06696939468383789
step: 200, loss: 0.18566705286502838
step: 210, loss: 0.04108394309878349
step: 220, loss: 0.31758418679237366
step: 230, loss: 0.18663586676120758
step: 240, loss: 0.07860952615737915
step: 250, loss: 0.08661098033189774
step: 260, loss: 0.32025450468063354
step: 270, loss: 0.05141540616750717
step: 280, loss: 0.10472384095191956
step: 290, loss: 0.1078319177031517
step: 300, loss: 0.10784848779439926
step: 310, loss: 0.23922163248062134
step: 320, loss: 0.11708678305149078
step: 330, loss: 0.0450480692088604
step: 340, loss: 0.11280544847249985
step: 350, loss: 0.04474456235766411
step: 360, loss: 0.0515446811914444
step: 370, loss: 0.15784992277622223
step: 380, loss: 0.015614660456776619
step: 390, loss: 0.3859425187110901
step: 400, loss: 0.23812979459762573
step: 410, loss: 0.06368624418973923
step: 420, loss: 0.19381043314933777
step: 430, loss: 0.15462428331375122
step: 440, loss: 0.029763052240014076
step: 450, loss: 0.0583718977868557
step: 460, loss: 0.03451903164386749

```

step: 470, loss: 0.07134488970041275
step: 480, loss: 0.12916772067546844
step: 490, loss: 0.2863258421421051
step: 500, loss: 0.05490024760365486
step: 510, loss: 0.15002501010894775
step: 520, loss: 0.03965567424893379
step: 530, loss: 0.020651385188102722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8894052044609666, f1=0.8921523634694815,
best_f1=0.8921523634694815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17673049867153168
step: 10, loss: 0.11850208789110184
step: 20, loss: 0.022296080365777016
step: 30, loss: 0.031363632529973984
step: 40, loss: 0.2073325514793396
step: 50, loss: 0.2408607304096222
step: 60, loss: 0.036921799182891846
step: 70, loss: 0.09085609018802643
step: 80, loss: 0.2643811106681824
step: 90, loss: 0.024047549813985825
step: 100, loss: 0.1541527509689331
step: 110, loss: 0.015019550919532776
step: 120, loss: 0.12280485033988953
step: 130, loss: 0.1362505406141281
step: 140, loss: 0.05599187687039375
step: 150, loss: 0.10231215506792068
step: 160, loss: 0.0807800143957138
step: 170, loss: 0.08637659251689911
step: 180, loss: 0.020576929673552513
step: 190, loss: 0.05276418477296829
step: 200, loss: 0.04831460490822792
step: 210, loss: 0.021045222878456116
step: 220, loss: 0.06639345735311508
step: 230, loss: 0.05346333608031273
step: 240, loss: 0.03130370005965233
step: 250, loss: 0.06608964502811432
step: 260, loss: 0.016177747398614883
step: 270, loss: 0.23710939288139343
step: 280, loss: 0.15686464309692383

```

```

step: 290, loss: 0.07056015729904175
step: 300, loss: 0.056118499487638474
step: 310, loss: 0.0306426789611578
step: 320, loss: 0.25500476360321045
step: 330, loss: 0.07010539621114731
step: 340, loss: 0.01535874791443348
step: 350, loss: 0.0015296153724193573
step: 360, loss: 0.10407102853059769
step: 370, loss: 0.12147407978773117
step: 380, loss: 0.1589396446943283
step: 390, loss: 0.06826496124267578
step: 400, loss: 0.0768948569893837
step: 410, loss: 0.060298334807157516
step: 420, loss: 0.04609429091215134
step: 430, loss: 0.038734227418899536
step: 440, loss: 0.029458610340952873
step: 450, loss: 0.04531741887331009
step: 460, loss: 0.03757837787270546
step: 470, loss: 0.05328746140003204
step: 480, loss: 0.13538099825382233
step: 490, loss: 0.03620292618870735
step: 500, loss: 0.3302943706512451
step: 510, loss: 0.025609754025936127
step: 520, loss: 0.05416293814778328
step: 530, loss: 0.018254157155752182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9130841121495327, f1=0.907906976744186,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0656229704618454
step: 10, loss: 0.014680192805826664
step: 20, loss: 0.18334539234638214
step: 30, loss: 0.06550410389900208
step: 40, loss: 0.026685478165745735
step: 50, loss: 0.06927380710840225
step: 60, loss: 0.030475322157144547
step: 70, loss: 0.03563292324542999
step: 80, loss: 0.0025986209511756897
step: 90, loss: 0.02408052794635296
step: 100, loss: 0.054339345544576645

```

```
step: 110, loss: 0.003651284845545888
step: 120, loss: 0.02822660654783249
step: 130, loss: 0.0004531680897343904
step: 140, loss: 0.17558646202087402
step: 150, loss: 0.06698139011859894
step: 160, loss: 0.02985456772148609
step: 170, loss: 0.14936232566833496
step: 180, loss: 0.06578250229358673
step: 190, loss: 0.0661688894033432
step: 200, loss: 0.04191194474697113
step: 210, loss: 0.020192669704556465
step: 220, loss: 0.16313078999519348
step: 230, loss: 0.0508696511387825
step: 240, loss: 0.01841559074819088
step: 250, loss: 0.05909213051199913
step: 260, loss: 0.024024898186326027
step: 270, loss: 0.04019007831811905
step: 280, loss: 0.1517118513584137
step: 290, loss: 0.034489382058382034
step: 300, loss: 0.044386912137269974
step: 310, loss: 0.03925461694598198
step: 320, loss: 0.017658766359090805
step: 330, loss: 0.012707439251244068
step: 340, loss: 0.00973082147538662
step: 350, loss: 0.010470293462276459
step: 360, loss: 0.05203299969434738
step: 370, loss: 0.029970955103635788
step: 380, loss: 0.05176863819360733
step: 390, loss: 0.10136513411998749
step: 400, loss: 0.013416064903140068
step: 410, loss: 0.012445506639778614
step: 420, loss: 0.09031267464160919
step: 430, loss: 0.01779058948159218
step: 440, loss: 0.023976091295480728
step: 450, loss: 0.03341711685061455
step: 460, loss: 0.029223766177892685
step: 470, loss: 0.018844278529286385
step: 480, loss: 0.013256553560495377
step: 490, loss: 0.01049773022532463
step: 500, loss: 0.04225611314177513
step: 510, loss: 0.007631917484104633
step: 520, loss: 0.12952029705047607
step: 530, loss: 0.07152814418077469
/usr/local/lib/python3.7/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=0.9057665260196905, f1=0.8975791433891994,
best_f1=0.907906976744186
```

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

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

```
step: 0, loss: 0.03353516384959221  
step: 10, loss: 0.006592847872525454  
step: 20, loss: 0.02204366959631443  
step: 30, loss: 0.02849036455154419  
step: 40, loss: 0.020877325907349586  
step: 50, loss: 0.0012942645698785782  
step: 60, loss: 0.003432227997109294  
step: 70, loss: 0.01887158863246441  
step: 80, loss: 0.031544364988803864  
step: 90, loss: 0.1687779724597931  
step: 100, loss: 0.021077051758766174  
step: 110, loss: 0.14530543982982635  
step: 120, loss: 0.08087769150733948  
step: 130, loss: 0.001786744687706232  
step: 140, loss: 0.003496959339827299  
step: 150, loss: 0.0032628236804157495  
step: 160, loss: 0.19998852908611298  
step: 170, loss: 0.006451488938182592  
step: 180, loss: 0.01291837077587843  
step: 190, loss: 0.021578652784228325  
step: 200, loss: 0.015064024366438389  
step: 210, loss: 0.014406724832952023  
step: 220, loss: 0.010567707009613514  
step: 230, loss: 0.041600003838539124  
step: 240, loss: 0.005482962355017662  
step: 250, loss: 0.04679363965988159  
step: 260, loss: 0.004476682282984257  
step: 270, loss: 0.02820878103375435  
step: 280, loss: 0.3219447135925293  
step: 290, loss: 0.1686958372592926  
step: 300, loss: 0.004173868335783482  
step: 310, loss: 0.021678932011127472  
step: 320, loss: 0.009919504635035992  
step: 330, loss: 0.07296358793973923  
step: 340, loss: 0.018259229138493538  
step: 350, loss: 0.007616249844431877  
step: 360, loss: 0.05933232977986336  
step: 370, loss: 0.006649344693869352  
step: 380, loss: 0.009567067958414555  
step: 390, loss: 0.004400509875267744  
step: 400, loss: 0.000887950649484992  
step: 410, loss: 0.0017772488063201308  
step: 420, loss: 0.07975087314844131  
step: 430, loss: 0.10394899547100067  
step: 440, loss: 0.0029089783784002066
```



```
step: 450, loss: 0.012085688300430775
step: 460, loss: 0.14197999238967896
step: 470, loss: 0.052806273102760315
step: 480, loss: 0.00837006140500307
step: 490, loss: 0.002976265037432313
step: 500, loss: 0.0027047665789723396
step: 510, loss: 0.018644364550709724
step: 520, loss: 0.027247266843914986
step: 530, loss: 0.0761466920375824
/usr/local/lib/python3.7/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=0.9043151969981239, f1=0.8886810102899906,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04649779573082924
step: 10, loss: 0.046793367713689804
step: 20, loss: 0.004402705933898687
step: 30, loss: 0.005431700497865677
step: 40, loss: 0.017087819054722786
step: 50, loss: 0.04996394366025925
step: 60, loss: 0.1064944639801979
step: 70, loss: 0.028600111603736877
step: 80, loss: 0.009443500079214573
step: 90, loss: 0.04450186714529991
step: 100, loss: 0.0008584714378230274
step: 110, loss: 0.015509175136685371
step: 120, loss: 0.01076600980013609
step: 130, loss: 0.0003414626116864383
step: 140, loss: 0.005138979759067297
step: 150, loss: 0.03876461461186409
step: 160, loss: 0.0010072062723338604
step: 170, loss: 0.011460518464446068
step: 180, loss: 0.0012928476789966226
step: 190, loss: 0.002069752663373947
step: 200, loss: 0.002599061466753483
step: 210, loss: 0.05518863722681999
step: 220, loss: 0.005342953838407993
step: 230, loss: 0.01210671104490757
step: 240, loss: 0.03010203130543232
step: 250, loss: 0.08481896668672562
step: 260, loss: 0.020281247794628143
step: 270, loss: 0.0029210352804511786
step: 280, loss: 0.030998336151242256
step: 290, loss: 0.20965369045734406
step: 300, loss: 0.006623988971114159
```

```

step: 310, loss: 0.014979755505919456
step: 320, loss: 0.1477387398481369
step: 330, loss: 0.007902687415480614
step: 340, loss: 0.004220015835016966
step: 350, loss: 0.000954215123783797
step: 360, loss: 0.002320877742022276
step: 370, loss: 0.02153163030743599
step: 380, loss: 0.0604398287832737
step: 390, loss: 0.00028877536533400416
step: 400, loss: 0.025101302191615105
step: 410, loss: 0.0055748093873262405
step: 420, loss: 0.004120148718357086
step: 430, loss: 0.0031333058141171932
step: 440, loss: 0.14039458334445953
step: 450, loss: 0.010162405669689178
step: 460, loss: 0.0015783135313540697
step: 470, loss: 0.007803116925060749
step: 480, loss: 0.0023016072809696198
step: 490, loss: 0.006762573961168528
step: 500, loss: 0.13018062710762024
step: 510, loss: 0.02139541693031788
step: 520, loss: 0.05483652651309967
step: 530, loss: 0.009183441288769245
/usr/local/lib/python3.7/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=0.9093484419263457, f1=0.9010367577756833,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007097164634615183
step: 10, loss: 0.00043430112418718636
step: 20, loss: 0.06318381428718567
step: 30, loss: 0.0009201960056088865
step: 40, loss: 0.08963903039693832
step: 50, loss: 0.010568267665803432
step: 60, loss: 0.00043033939437009394
step: 70, loss: 0.00045018261880613863
step: 80, loss: 0.0015712975291535258
step: 90, loss: 0.0003514218842610717
step: 100, loss: 0.008873219601809978
step: 110, loss: 0.00018667406402528286
step: 120, loss: 8.699952741153538e-05
step: 130, loss: 0.0005848250002600253
step: 140, loss: 0.006524394731968641
step: 150, loss: 0.0013045478845015168
step: 160, loss: 0.0024931614752858877

```

```

step: 170, loss: 0.0030491091310977936
step: 180, loss: 0.0007763973553664982
step: 190, loss: 0.00368082569912076
step: 200, loss: 0.0019407665822654963
step: 210, loss: 0.014341533184051514
step: 220, loss: 0.009833596646785736
step: 230, loss: 0.0035073780454695225
step: 240, loss: 0.03674260526895523
step: 250, loss: 0.0022654852364212275
step: 260, loss: 0.0002771934377960861
step: 270, loss: 0.01834707520902157
step: 280, loss: 0.08340664952993393
step: 290, loss: 0.0039887940511107445
step: 300, loss: 0.005706240888684988
step: 310, loss: 0.0010395270073786378
step: 320, loss: 0.0010213463101536036
step: 330, loss: 0.0007274713134393096
step: 340, loss: 0.059652335941791534
step: 350, loss: 0.0077919624745845795
step: 360, loss: 0.006135732866823673
step: 370, loss: 0.02480163984000683
step: 380, loss: 0.00012867318582721055
step: 390, loss: 0.009060335345566273
step: 400, loss: 0.05479762330651283
step: 410, loss: 0.0005236480501480401
step: 420, loss: 0.007391876541078091
step: 430, loss: 0.00011212180834263563
step: 440, loss: 0.006963890977203846
step: 450, loss: 0.0014583806041628122
step: 460, loss: 0.0015608163084834814
step: 470, loss: 0.0446413978934288
step: 480, loss: 0.006107311230152845
step: 490, loss: 0.010246988385915756
step: 500, loss: 0.02640393190085888
step: 510, loss: 0.008018814958631992
step: 520, loss: 0.005203367210924625
step: 530, loss: 0.018189148977398872
/usr/local/lib/python3.7/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=0.8698254364089776, f1=0.8650246305418717,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007003589998930693
step: 10, loss: 0.06703511625528336
step: 20, loss: 0.009680020622909069

```

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

step: 30, loss: 0.010342423804104328
step: 40, loss: 0.06456290930509567
step: 50, loss: 0.03799606114625931
step: 60, loss: 0.0009591329144313931
step: 70, loss: 0.04144798591732979
step: 80, loss: 0.0022351485677063465
step: 90, loss: 0.0003032982349395752
step: 100, loss: 0.0008109543123282492
step: 110, loss: 0.022038482129573822
step: 120, loss: 0.000134759015054442
step: 130, loss: 0.00048080930719152093
step: 140, loss: 0.00012479812721721828
step: 150, loss: 0.0028075582813471556
step: 160, loss: 0.014474156312644482
step: 170, loss: 0.008093430660665035
step: 180, loss: 0.0008092163479886949
step: 190, loss: 0.0006828279001638293
step: 200, loss: 0.0023677183780819178
step: 210, loss: 0.04986673220992088
step: 220, loss: 0.00045798346400260925
step: 230, loss: 0.000415063404943794
step: 240, loss: 0.17604242265224457
step: 250, loss: 0.001288966042920947
step: 260, loss: 0.00043163340887986124
step: 270, loss: 0.0012410138733685017
step: 280, loss: 0.004031619522720575
step: 290, loss: 0.08485060185194016
step: 300, loss: 0.0012959556188434362
step: 310, loss: 0.000780905713327229
step: 320, loss: 0.00022586261911783367
step: 330, loss: 0.02612793631851673
step: 340, loss: 0.007332670968025923
step: 350, loss: 0.001595870591700077
step: 360, loss: 0.00013253815995994955
step: 370, loss: 0.00025721630663610995
step: 380, loss: 0.00025154545437544584
step: 390, loss: 0.00015771928883623332
step: 400, loss: 0.00012645173410419375
step: 410, loss: 0.00577740790322423
step: 420, loss: 0.02110503613948822
step: 430, loss: 5.73932848055847e-05
step: 440, loss: 0.0008775625028647482
step: 450, loss: 0.0004149508604314178
step: 460, loss: 7.405014184769243e-05
step: 470, loss: 0.0006372568313963711
step: 480, loss: 0.05368507653474808
step: 490, loss: 0.0018681437941268086

```
step: 500, loss: 0.03547088801860809
step: 510, loss: 0.0010748793138191104
step: 520, loss: 0.0015564317582175136
step: 530, loss: 0.001956926891580224
/usr/local/lib/python3.7/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=0.9038112522686025, f1=0.902461257976299,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03744813799858093
step: 10, loss: 0.005273173563182354
step: 20, loss: 0.002738777780905366
step: 30, loss: 0.08842939138412476
step: 40, loss: 0.00033845490543171763
step: 50, loss: 0.06851015239953995
step: 60, loss: 0.0010993318865075707
step: 70, loss: 0.0012232671724632382
step: 80, loss: 0.0014289767714217305
step: 90, loss: 0.030217742547392845
step: 100, loss: 0.0010320758447051048
step: 110, loss: 0.0005751218996010721
step: 120, loss: 0.0030361551325768232
step: 130, loss: 0.0001974858168978244
step: 140, loss: 0.000318357371725142
step: 150, loss: 0.0032966313883662224
step: 160, loss: 0.0005788921262137592
step: 170, loss: 0.0019726769533008337
step: 180, loss: 0.0009451636578887701
step: 190, loss: 0.011135276407003403
step: 200, loss: 0.003057154593989253
step: 210, loss: 0.001963770017027855
step: 220, loss: 0.00040565821109339595
step: 230, loss: 0.0015448424965143204
step: 240, loss: 0.005641024559736252
step: 250, loss: 7.303973688976839e-05
step: 260, loss: 4.07899497076869e-05
step: 270, loss: 0.007535536307841539
step: 280, loss: 0.026853235438466072
step: 290, loss: 0.0002465834841132164
step: 300, loss: 0.10509040951728821
step: 310, loss: 0.0004684123850893229
step: 320, loss: 0.08630082756280899
step: 330, loss: 0.01103284116834402
step: 340, loss: 0.0003467428614385426
step: 350, loss: 0.03482377156615257
```

```
step: 360, loss: 0.00361900357529521
step: 370, loss: 0.04599525406956673
step: 380, loss: 0.003625014331191778
step: 390, loss: 0.07388181239366531
step: 400, loss: 0.0004451691056601703
step: 410, loss: 0.00016514411254320294
step: 420, loss: 0.00036810923484154046
step: 430, loss: 0.04089663177728653
step: 440, loss: 0.015799451619386673
step: 450, loss: 0.00021424159058369696
step: 460, loss: 0.0008006665157154202
step: 470, loss: 0.004865577444434166
step: 480, loss: 0.0002241030742879957
step: 490, loss: 0.005466548725962639
step: 500, loss: 0.0013514168094843626
step: 510, loss: 0.001756887068040669
step: 520, loss: 0.0024218924809247255
step: 530, loss: 0.00014416061458177865
/usr/local/lib/python3.7/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=0.9067041725269573, f1=0.9010270774976656,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004598097875714302
step: 10, loss: 0.0006197154289111495
step: 20, loss: 0.0008747333195060492
step: 30, loss: 0.006807386875152588
step: 40, loss: 5.5690052249701694e-05
step: 50, loss: 0.002492243889719248
step: 60, loss: 5.561108628171496e-05
step: 70, loss: 0.0004697063995990902
step: 80, loss: 0.019245782867074013
step: 90, loss: 0.0020466449204832315
step: 100, loss: 0.029938848689198494
step: 110, loss: 0.0083073815330863
step: 120, loss: 0.00042955257231369615
step: 130, loss: 0.0002942630962934345
step: 140, loss: 0.0002682960475794971
step: 150, loss: 0.000542782130651176
step: 160, loss: 0.0012348211603239179
step: 170, loss: 0.07632629573345184
step: 180, loss: 0.00041341324686072767
step: 190, loss: 5.679806417902e-05
step: 200, loss: 0.0002920165716204792
step: 210, loss: 7.076412293827161e-05
```

```

step: 220, loss: 0.07116734981536865
step: 230, loss: 0.0007179688545875251
step: 240, loss: 0.00044546640128828585
step: 250, loss: 0.0009384016739204526
step: 260, loss: 0.0008973692310974002
step: 270, loss: 0.0009820314589887857
step: 280, loss: 0.0009578831377439201
step: 290, loss: 0.026141300797462463
step: 300, loss: 0.0001563061523484066
step: 310, loss: 0.009391573257744312
step: 320, loss: 0.0002408196305623278
step: 330, loss: 0.12340768426656723
step: 340, loss: 0.0013694643275812268
step: 350, loss: 0.00017412954184692353
step: 360, loss: 0.0015333020128309727
step: 370, loss: 0.00026409272686578333
step: 380, loss: 0.0001174499120679684
step: 390, loss: 0.000167490667081438
step: 400, loss: 0.0003088793600909412
step: 410, loss: 3.646173718152568e-05
step: 420, loss: 6.853398372186348e-05
step: 430, loss: 8.4783008787781e-05
step: 440, loss: 0.00023741314362268895
step: 450, loss: 0.00011421633826103061
step: 460, loss: 9.337843948742375e-05
step: 470, loss: 3.253195245633833e-05
step: 480, loss: 4.5609151129610837e-05
step: 490, loss: 0.013635894283652306
step: 500, loss: 7.539673242717981e-05
step: 510, loss: 9.649269486544654e-05
step: 520, loss: 4.0987513784784824e-05
step: 530, loss: 0.0006004719762131572
/usr/local/lib/python3.7/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=0.906468124709167, f1=0.9023041474654377,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.715479634702206e-05
step: 10, loss: 0.0007243562722578645
step: 20, loss: 0.07000479102134705
step: 30, loss: 3.758272578124888e-05
step: 40, loss: 0.0005423637921921909
step: 50, loss: 4.4715347030432895e-05
step: 60, loss: 5.0842932978412136e-05
step: 70, loss: 0.00029455136973410845

```

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

step: 80, loss: 2.9492286557797343e-05
step: 90, loss: 0.00017994949303101748
step: 100, loss: 0.00097120605641976
step: 110, loss: 4.21455260948278e-05
step: 120, loss: 0.00032321581966243684
step: 130, loss: 0.00909137912094593
step: 140, loss: 0.01764713227748871
step: 150, loss: 0.01067001186311245
step: 160, loss: 5.465357389766723e-05
step: 170, loss: 0.00013220588152762502
step: 180, loss: 0.00029195594834163785
step: 190, loss: 4.315561818657443e-05
step: 200, loss: 0.00012278323993086815
step: 210, loss: 0.00013621142716147006
step: 220, loss: 0.00019125152903143317
step: 230, loss: 0.00011444420670159161
step: 240, loss: 0.00019515692838467658
step: 250, loss: 0.0002915819932240993
step: 260, loss: 0.003880760632455349
step: 270, loss: 8.148855704348534e-05
step: 280, loss: 6.190451676957309e-05
step: 290, loss: 0.00303435605019331
step: 300, loss: 3.5034474421991035e-05
step: 310, loss: 0.0006713205366395414
step: 320, loss: 0.008919217623770237
step: 330, loss: 0.006242799106985331
step: 340, loss: 0.009267936460673809
step: 350, loss: 3.8011356082279235e-05
step: 360, loss: 5.0566835852805525e-05
step: 370, loss: 0.00014800456119701266
step: 380, loss: 8.847173739923164e-05
step: 390, loss: 0.009819644503295422
step: 400, loss: 3.24834072671365e-05
step: 410, loss: 5.570549183175899e-05
step: 420, loss: 8.324395457748324e-05
step: 430, loss: 2.5957011530408636e-05
step: 440, loss: 0.0004991673631593585
step: 450, loss: 2.58118689089315e-05
step: 460, loss: 9.555602446198463e-05
step: 470, loss: 8.575192623538896e-05
step: 480, loss: 0.00014165914035402238
step: 490, loss: 0.0002081673446809873
step: 500, loss: 0.018253013491630554
step: 510, loss: 4.162975528743118e-05
step: 520, loss: 0.00014875653141643852
step: 530, loss: 0.00023797660833224654

/usr/local/lib/python3.7/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=0.9064885496183206, f1=0.8964200477326969,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.081291965325363e-05
step: 10, loss: 0.0025651496835052967
step: 20, loss: 6.741225661244243e-05
step: 30, loss: 0.00025323216686956584
step: 40, loss: 7.429214019794017e-05
step: 50, loss: 0.0015767216682434082
step: 60, loss: 0.01819775253534317
step: 70, loss: 0.02727392502129078
step: 80, loss: 4.019670086563565e-05
step: 90, loss: 0.00013868566020391881
step: 100, loss: 0.004339035600423813
step: 110, loss: 0.0003929919912479818
step: 120, loss: 2.825921001203824e-05
step: 130, loss: 6.398349069058895e-05
step: 140, loss: 0.0003628600970841944
step: 150, loss: 4.1838138713501394e-05
step: 160, loss: 3.146612652926706e-05
step: 170, loss: 4.0790171624394134e-05
step: 180, loss: 0.00027479504933580756
step: 190, loss: 0.0032147751189768314
step: 200, loss: 0.0004104043764527887
step: 210, loss: 0.005016573704779148
step: 220, loss: 9.030909859575331e-05
step: 230, loss: 2.411306013527792e-05
step: 240, loss: 0.013562125153839588
step: 250, loss: 8.066402369877324e-05
step: 260, loss: 5.921621777815744e-05
step: 270, loss: 0.026852751150727272
step: 280, loss: 0.00048080080887302756
step: 290, loss: 0.0004969134461134672
step: 300, loss: 0.00016959263302851468
step: 310, loss: 4.2734282033052295e-05
step: 320, loss: 0.0001144821144407615
step: 330, loss: 0.0008062546257860959
step: 340, loss: 0.0003733958292286843
step: 350, loss: 7.142642425606027e-05
step: 360, loss: 3.3792945032473654e-05
step: 370, loss: 5.190061710891314e-05
step: 380, loss: 2.9820130293956026e-05
step: 390, loss: 2.7536523703020066e-05
step: 400, loss: 3.319778261356987e-05
```

```

step: 410, loss: 8.280971087515354e-05
step: 420, loss: 0.00013244197180029005
step: 430, loss: 3.171979551552795e-05
step: 440, loss: 4.84401207359042e-05
step: 450, loss: 0.0008212481043301523
step: 460, loss: 0.014873804524540901
step: 470, loss: 0.00013361609308049083
step: 480, loss: 0.0003047179488930851
step: 490, loss: 5.201433668844402e-05
step: 500, loss: 0.00011745494703063741
step: 510, loss: 3.69892550224904e-05
step: 520, loss: 0.039620768278837204
step: 530, loss: 0.001577003626152873
/usr/local/lib/python3.7/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=0.9038010323791646, f1=0.9014479215319944,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.0467814212897792e-05
step: 10, loss: 2.6452609745319933e-05
step: 20, loss: 0.007990064099431038
step: 30, loss: 6.109716196078807e-05
step: 40, loss: 3.1458686862606555e-05
step: 50, loss: 0.006931259296834469
step: 60, loss: 0.00012784203863702714
step: 70, loss: 0.0001162027328973636
step: 80, loss: 0.00029480073135346174
step: 90, loss: 0.0003251723537687212
step: 100, loss: 0.11836361885070801
step: 110, loss: 0.00021221187489572912
step: 120, loss: 2.8742697395500727e-05
step: 130, loss: 7.879212353145704e-05
step: 140, loss: 5.0269434723304585e-05
step: 150, loss: 0.00012331240577623248
step: 160, loss: 4.843120768782683e-05
step: 170, loss: 0.0029972626361995935
step: 180, loss: 2.5826828277786262e-05
step: 190, loss: 3.6323192034615204e-05
step: 200, loss: 8.44937821966596e-05
step: 210, loss: 4.173124034423381e-05
step: 220, loss: 0.0001884972007246688
step: 230, loss: 3.0177348889992572e-05
step: 240, loss: 0.00022562970116268843
step: 250, loss: 2.4552737158956006e-05
step: 260, loss: 5.400657028076239e-05

```

```

step: 270, loss: 5.684009374817833e-05
step: 280, loss: 5.13661143486388e-05
step: 290, loss: 0.002952966606244445
step: 300, loss: 0.0002066579763777554
step: 310, loss: 0.00014484288112726063
step: 320, loss: 8.834216714603826e-05
step: 330, loss: 0.0002153700334019959
step: 340, loss: 0.0001708849158603698
step: 350, loss: 0.0030172253027558327
step: 360, loss: 8.542503928765655e-05
step: 370, loss: 4.867528332397342e-05
step: 380, loss: 5.763319495599717e-05
step: 390, loss: 4.0987513784784824e-05
step: 400, loss: 0.020445657894015312
step: 410, loss: 8.498416718794033e-05
step: 420, loss: 0.0003158478648401797
step: 430, loss: 0.00798824056982994
step: 440, loss: 0.0001608507300261408
step: 450, loss: 5.975125532131642e-05
step: 460, loss: 3.3652653655735776e-05
step: 470, loss: 6.693789327982813e-05
step: 480, loss: 0.00017594634846318513
step: 490, loss: 7.086930418154225e-05
step: 500, loss: 0.00032459208159707487
step: 510, loss: 8.133615483529866e-05
step: 520, loss: 5.798911661258899e-05
step: 530, loss: 3.438662315602414e-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=0.9115128449096099, f1=0.8944866920152091,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.4567215556744486e-05
step: 10, loss: 0.0003799335681833327
step: 20, loss: 2.839342960214708e-05
step: 30, loss: 6.285351264523342e-05
step: 40, loss: 4.8759273340692744e-05
step: 50, loss: 0.0005485267611220479
step: 60, loss: 0.0005747061804868281
step: 70, loss: 5.645591954817064e-05
step: 80, loss: 4.921135769109242e-05
step: 90, loss: 0.0001725159672787413
step: 100, loss: 0.00022950171842239797
step: 110, loss: 1.8082249880535528e-05
step: 120, loss: 3.9531627407995984e-05

```

```
step: 130, loss: 3.078098961850628e-05
step: 140, loss: 3.454981560935266e-05
step: 150, loss: 4.4368633098201826e-05
step: 160, loss: 4.039308623760007e-05
step: 170, loss: 4.911808355245739e-05
step: 180, loss: 7.671162893529981e-05
step: 190, loss: 0.00021668720000889152
step: 200, loss: 0.0001701967412373051
step: 210, loss: 0.0001848704123403877
step: 220, loss: 3.728749288711697e-05
step: 230, loss: 4.10340289818123e-05
step: 240, loss: 0.0007354443077929318
step: 250, loss: 1.8849685147870332e-05
step: 260, loss: 2.4317942006746307e-05
step: 270, loss: 2.4508015485480428e-05
step: 280, loss: 2.6708594305091538e-05
step: 290, loss: 4.4980701204622164e-05
step: 300, loss: 5.4843927500769496e-05
step: 310, loss: 4.811702820006758e-05
step: 320, loss: 0.0008109326590783894
step: 330, loss: 2.4727856725803576e-05
step: 340, loss: 6.662256055278704e-05
step: 350, loss: 0.030849333852529526
step: 360, loss: 4.4124357373220846e-05
step: 370, loss: 5.712005804525688e-05
step: 380, loss: 3.5886430850951e-05
step: 390, loss: 5.118290937389247e-05
step: 400, loss: 1.903964948724024e-05
step: 410, loss: 3.2110267056850716e-05
step: 420, loss: 2.2656724468106404e-05
step: 430, loss: 7.171651668613777e-05
step: 440, loss: 0.0001236946991411969
step: 450, loss: 0.002919713268056512
step: 460, loss: 0.00012283850810490549
step: 470, loss: 8.668781811138615e-05
step: 480, loss: 1.9326527763041668e-05
step: 490, loss: 4.293804886401631e-05
step: 500, loss: 3.316570655442774e-05
step: 510, loss: 0.00010173898044740781
step: 520, loss: 0.0023510130122303963
step: 530, loss: 1.9669223547680303e-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=0.9096958174904942, f1=0.8971076339497392,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.6953922467073426e-05
step: 10, loss: 1.9147502825944684e-05
step: 20, loss: 2.304415284015704e-05
step: 30, loss: 4.051625364809297e-05
step: 40, loss: 4.877046012552455e-05
step: 50, loss: 0.0001917577610583976
step: 60, loss: 0.00012900863657705486
step: 70, loss: 1.9158578652422875e-05
step: 80, loss: 3.7189314753049985e-05
step: 90, loss: 2.671679249033332e-05
step: 100, loss: 3.083585397689603e-05
step: 110, loss: 0.08584386110305786
step: 120, loss: 2.0868708816124126e-05
step: 130, loss: 9.911884262692183e-05
step: 140, loss: 2.271963057864923e-05
step: 150, loss: 0.0005925916484557092
step: 160, loss: 2.7585077987168916e-05
step: 170, loss: 3.974654828198254e-05
step: 180, loss: 2.754727211140562e-05
step: 190, loss: 6.660680810455233e-05
step: 200, loss: 3.859897697111592e-05
step: 210, loss: 9.164456423604861e-05
step: 220, loss: 1.9449427782092243e-05
step: 230, loss: 4.908811752102338e-05
step: 240, loss: 2.6895515475189313e-05
step: 250, loss: 3.942466719308868e-05
step: 260, loss: 2.3677150238654576e-05
step: 270, loss: 2.851957469829358e-05
step: 280, loss: 0.000369972171029076
step: 290, loss: 0.003451727097854018
step: 300, loss: 2.111822323058732e-05
step: 310, loss: 5.832535680383444e-05
step: 320, loss: 7.117716450011358e-05
step: 330, loss: 0.002169401617720723
step: 340, loss: 0.00012538822193164378
step: 350, loss: 7.612331683048978e-05
step: 360, loss: 5.020895696361549e-05
step: 370, loss: 5.242408951744437e-05
step: 380, loss: 7.180296961450949e-05
step: 390, loss: 0.006864104885607958
step: 400, loss: 8.865356358001009e-05
step: 410, loss: 1.678594526310917e-05
step: 420, loss: 0.021187396720051765
step: 430, loss: 8.476684160996228e-05
step: 440, loss: 0.00022617136710323393
step: 450, loss: 0.0004816438886336982
step: 460, loss: 0.00012659186904784292
```

```

step: 470, loss: 0.0003476636193227023
step: 480, loss: 2.5658533559180796e-05
step: 490, loss: 0.03165891766548157
step: 500, loss: 1.6409710951847956e-05
step: 510, loss: 0.0006326469010673463
step: 520, loss: 4.682780854636803e-05
step: 530, loss: 2.093206421704963e-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=0.9097744360902255, f1=0.8990136214185063,
best_f1=0.907906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2123955204733647e-05
step: 10, loss: 0.0002543407026678324
step: 20, loss: 6.462400779128075e-05
step: 30, loss: 2.2883945348439738e-05
step: 40, loss: 0.009490042924880981
step: 50, loss: 6.173301517264917e-05
step: 60, loss: 1.7747064703144133e-05
step: 70, loss: 3.151281998725608e-05
step: 80, loss: 0.0005914279608987272
step: 90, loss: 0.002680284669622779
step: 100, loss: 6.23565647401847e-05
step: 110, loss: 0.0003040523442905396
step: 120, loss: 0.00155977054964751
step: 130, loss: 0.0031226053833961487
step: 140, loss: 1.98219386220444e-05
step: 150, loss: 3.845838364213705e-05
step: 160, loss: 0.0007427555974572897
step: 170, loss: 6.818433030275628e-05
step: 180, loss: 1.5750349120935425e-05
step: 190, loss: 8.178925781976432e-05
step: 200, loss: 4.178916788077913e-05
step: 210, loss: 6.0428286815294996e-05
step: 220, loss: 8.169822103809565e-05
step: 230, loss: 5.485356814460829e-05
step: 240, loss: 0.0004334877012297511
step: 250, loss: 2.766982106550131e-05
step: 260, loss: 5.469026291393675e-05
step: 270, loss: 4.5084652811056e-05
step: 280, loss: 2.5591840312699787e-05
step: 290, loss: 0.0003985149960499257
step: 300, loss: 2.1103191102156416e-05
step: 310, loss: 0.0001188963942695409
step: 320, loss: 0.0002859432715922594

```

```

step: 330, loss: 2.0887258870061487e-05
step: 340, loss: 1.8875773093895987e-05
step: 350, loss: 1.3034687071922235e-05
step: 360, loss: 0.0005008774460293353
step: 370, loss: 1.418576266587479e-05
step: 380, loss: 7.598836964461952e-05
step: 390, loss: 1.8641077986103483e-05
step: 400, loss: 4.5020027755526826e-05
step: 410, loss: 8.097948739305139e-05
step: 420, loss: 1.9192379113519564e-05
step: 430, loss: 0.00011925349826924503
step: 440, loss: 2.490665610821452e-05
step: 450, loss: 1.680076093180105e-05
step: 460, loss: 2.753262197074946e-05
step: 470, loss: 8.78496648510918e-05
step: 480, loss: 1.3030932677793317e-05
step: 490, loss: 1.9240806068410166e-05
step: 500, loss: 7.150138117140159e-05
step: 510, loss: 7.829145761206746e-05
step: 520, loss: 2.3684520783717744e-05
step: 530, loss: 1.9337672711117193e-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=0.9110486891385768, f1=0.90014064697609,
best_f1=0.907906976744186
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:31, 181.39it/s]
load_f1 = 0.9111826967326276
real_f1 = 0.9099264705882353
0it [00:00, ?it/s]/usr/local/lib/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, 180.28it/s]

```

3.1.4 Amazon-Google - Running the matcher

```

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

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

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/Amazon-Google \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \
    --lm bert-base-uncased \

```



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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

```
step: 0, loss: 0.5483188629150391  
step: 10, loss: 0.39112716913223267  
step: 20, loss: 0.3592701256275177  
step: 30, loss: 0.310160756111145  
step: 40, loss: 0.19110792875289917  
step: 50, loss: 0.4393254220485687  
step: 60, loss: 0.28172147274017334
```

```

step: 70, loss: 0.2521669566631317
step: 80, loss: 0.18054641783237457
step: 90, loss: 0.44780194759368896
step: 100, loss: 0.49234819412231445
step: 110, loss: 0.275777131319046
step: 120, loss: 0.28153789043426514
step: 130, loss: 0.3191806674003601
step: 140, loss: 0.23763032257556915
step: 150, loss: 0.246140718460083
step: 160, loss: 0.2969309091567993
step: 170, loss: 0.28456273674964905
step: 180, loss: 0.1342092901468277
step: 190, loss: 0.23065736889839172
step: 200, loss: 0.3565085232257843
step: 210, loss: 0.28061747550964355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.42689075630252094, f1=0.4825291181364393,
best_f1=0.4825291181364393
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17592360079288483
step: 10, loss: 0.32136163115501404
step: 20, loss: 0.26048094034194946
step: 30, loss: 0.42373278737068176
step: 40, loss: 0.21594516932964325
step: 50, loss: 0.14349104464054108
step: 60, loss: 0.45761168003082275
step: 70, loss: 0.10866077989339828
step: 80, loss: 0.2643962502479553
step: 90, loss: 0.17249959707260132
step: 100, loss: 0.03747202455997467
step: 110, loss: 0.19175437092781067
step: 120, loss: 0.1478605568408966
step: 130, loss: 0.0662771463394165
step: 140, loss: 0.2600002586841583
step: 150, loss: 0.300148606300354
step: 160, loss: 0.18532179296016693
step: 170, loss: 0.19148415327072144
step: 180, loss: 0.23884481191635132
step: 190, loss: 0.28407397866249084
step: 200, loss: 0.060276489704847336

```

```

step: 210, loss: 0.1669456660747528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4927536231884058, f1=0.5326460481099656,
best_f1=0.5326460481099656
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061123743653297424
step: 10, loss: 0.17951901257038116
step: 20, loss: 0.25437796115875244
step: 30, loss: 0.21894362568855286
step: 40, loss: 0.05147017911076546
step: 50, loss: 0.12928655743598938
step: 60, loss: 0.20532634854316711
step: 70, loss: 0.2488166242837906
step: 80, loss: 0.1493535190820694
step: 90, loss: 0.0677238181233406
step: 100, loss: 0.37492576241493225
step: 110, loss: 0.23653928935527802
step: 120, loss: 0.14444813132286072
step: 130, loss: 0.22639767825603485
step: 140, loss: 0.060752883553504944
step: 150, loss: 0.3008284270763397
step: 160, loss: 0.06103229522705078
step: 170, loss: 0.12858234345912933
step: 180, loss: 0.06492402404546738
step: 190, loss: 0.3098655045032501
step: 200, loss: 0.16778743267059326
step: 210, loss: 0.1633566915988922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5034013605442177, f1=0.5495652173913044,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2457180917263031

```

```

step: 10, loss: 0.16976352035999298
step: 20, loss: 0.06281518936157227
step: 30, loss: 0.034646663814783096
step: 40, loss: 0.016136042773723602
step: 50, loss: 0.19907675683498383
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.11328615248203278
step: 70, loss: 0.16774338483810425
step: 80, loss: 0.12698210775852203
step: 90, loss: 0.043678976595401764
step: 100, loss: 0.3043454885482788
step: 110, loss: 0.13029366731643677
step: 120, loss: 0.286465585231781
step: 130, loss: 0.037940721958875656
step: 140, loss: 0.2298945188522339
step: 150, loss: 0.10278904438018799
step: 160, loss: 0.05831747129559517
step: 170, loss: 0.16081523895263672
step: 180, loss: 0.4968874156475067
step: 190, loss: 0.19321541488170624
step: 200, loss: 0.18598762154579163
step: 210, loss: 0.2612646818161011
/usr/local/lib/python3.7/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=0.48704663212435234, f1=0.5390070921985817,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10974933207035065
step: 10, loss: 0.08107291907072067
step: 20, loss: 0.19366610050201416
step: 30, loss: 0.07422488927841187
step: 40, loss: 0.10229698568582535
step: 50, loss: 0.1882661134004593
step: 60, loss: 0.11225685477256775
step: 70, loss: 0.024206720292568207
step: 80, loss: 0.03646409511566162
step: 90, loss: 0.05686299502849579
step: 100, loss: 0.014499866403639317
step: 110, loss: 0.35422396659851074
step: 120, loss: 0.12010897696018219
step: 130, loss: 0.11459162086248398
step: 140, loss: 0.18662485480308533
step: 150, loss: 0.1325116753578186
step: 160, loss: 0.18872500956058502
step: 170, loss: 0.15987247228622437

```

```

step: 180, loss: 0.10287914425134659
step: 190, loss: 0.06016924977302551
step: 200, loss: 0.12206082791090012
step: 210, loss: 0.03155854344367981
/usr/local/lib/python3.7/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=0.49056603773584906, f1=0.5401709401709401,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027360746636986732
step: 10, loss: 0.12549595534801483
step: 20, loss: 0.020814865827560425
step: 30, loss: 0.001443990389816463
step: 40, loss: 0.053083766251802444
step: 50, loss: 0.004140131641179323
step: 60, loss: 0.14695769548416138
step: 70, loss: 0.008929332718253136
step: 80, loss: 0.028452571481466293
step: 90, loss: 0.16339422762393951
step: 100, loss: 0.004145897924900055
step: 110, loss: 0.010586568154394627
step: 120, loss: 0.03988558053970337
step: 130, loss: 0.2025052160024643
step: 140, loss: 0.09680500626564026
step: 150, loss: 0.028014488518238068
step: 160, loss: 0.05320383235812187
step: 170, loss: 0.11887655407190323
step: 180, loss: 0.08956947177648544
step: 190, loss: 0.07059872150421143
step: 200, loss: 0.09273654222488403
step: 210, loss: 0.04989178851246834
/usr/local/lib/python3.7/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=0.4564459930313589, f1=0.5159010600706714,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019381260499358177
step: 10, loss: 0.08939484506845474
step: 20, loss: 0.0675668865442276
step: 30, loss: 0.09879879653453827
step: 40, loss: 0.013561815023422241
step: 50, loss: 0.2024102509021759

```

```

step: 60, loss: 0.025719575583934784
step: 70, loss: 0.012758774682879448
step: 80, loss: 0.09855744987726212
step: 90, loss: 0.11985883116722107
step: 100, loss: 0.005115697160363197
step: 110, loss: 0.17156538367271423
step: 120, loss: 0.06943763047456741
step: 130, loss: 0.038067858666181564
step: 140, loss: 0.03129391372203827
step: 150, loss: 0.02262812852859497
step: 160, loss: 0.004404081031680107
step: 170, loss: 0.0037812944501638412
step: 180, loss: 0.017844995483756065
step: 190, loss: 0.004538544919341803
step: 200, loss: 0.0025854504201561213
step: 210, loss: 0.01669134385883808
/usr/local/lib/python3.7/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=0.43340857787810383, f1=0.5023474178403756,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008050496689975262
step: 10, loss: 0.10780119895935059
step: 20, loss: 0.011305198073387146
step: 30, loss: 0.0030377504881471395
step: 40, loss: 0.018894486129283905
step: 50, loss: 0.006257252302020788
step: 60, loss: 0.04925892874598503
step: 70, loss: 0.010714668780565262
step: 80, loss: 0.07473457604646683
step: 90, loss: 0.018512947484850883
step: 100, loss: 0.019160529598593712
step: 110, loss: 0.022938812151551247
step: 120, loss: 0.0038844896480441093
step: 130, loss: 0.003034529509022832
step: 140, loss: 0.004722025711089373
step: 150, loss: 0.014396683312952518
step: 160, loss: 0.06853161752223969
step: 170, loss: 0.009695881977677345
step: 180, loss: 0.17145639657974243
step: 190, loss: 0.01258703414350748
step: 200, loss: 0.01770460419356823
step: 210, loss: 0.07534687221050262
/usr/local/lib/python3.7/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=0.4424379232505643, f1=0.488262910798122,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021899497136473656
step: 10, loss: 0.01603076048195362
step: 20, loss: 0.015169416554272175
step: 30, loss: 0.0067380876280367374
step: 40, loss: 0.050860460847616196
step: 50, loss: 0.011247918009757996
step: 60, loss: 0.05460711941123009
step: 70, loss: 0.008504031226038933
step: 80, loss: 0.002577386796474457
step: 90, loss: 0.0010394009295850992
step: 100, loss: 0.005187541246414185
step: 110, loss: 0.007695329375565052
step: 120, loss: 0.00045905227307230234
step: 130, loss: 0.020816057920455933
step: 140, loss: 0.020741119980812073
step: 150, loss: 0.08867024630308151
step: 160, loss: 0.0011419201036915183
step: 170, loss: 0.001213323324918747
step: 180, loss: 0.07081428915262222
step: 190, loss: 0.0008632439421489835
step: 200, loss: 0.08108383417129517
step: 210, loss: 0.0029881037771701813
/usr/local/lib/python3.7/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=0.44921875, f1=0.5096525096525096, best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013618886470794678
step: 10, loss: 0.004422257654368877
step: 20, loss: 0.008458462543785572
step: 30, loss: 0.01872064732015133
step: 40, loss: 0.005114738363772631
step: 50, loss: 0.004912805277854204
step: 60, loss: 0.005040478892624378
step: 70, loss: 0.02472073957324028
step: 80, loss: 0.00558092538267374
step: 90, loss: 0.03744924068450928
step: 100, loss: 0.025426382198929787
step: 110, loss: 0.0026574430521577597
step: 120, loss: 0.07138659060001373

```

```

step: 130, loss: 0.054908156394958496
step: 140, loss: 0.0007944648386910558
step: 150, loss: 0.02184680849313736
step: 160, loss: 0.005981167778372765
step: 170, loss: 0.006083168089389801
step: 180, loss: 0.005000079050660133
step: 190, loss: 0.0010193443158641458
step: 200, loss: 0.004860476590692997
step: 210, loss: 0.008428463712334633
/usr/local/lib/python3.7/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=0.4247787610619469, f1=0.5178571428571429,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003962431102991104
step: 10, loss: 0.00786309503018856
step: 20, loss: 0.0681886151432991
step: 30, loss: 0.0026255433913320303
step: 40, loss: 0.003367218654602766
step: 50, loss: 0.002372868126258254
step: 60, loss: 0.0015501207672059536
step: 70, loss: 0.05130873993039131
step: 80, loss: 0.003237874014303088
step: 90, loss: 0.0027370182797312737
step: 100, loss: 0.09913662075996399
step: 110, loss: 0.0011793682351708412
step: 120, loss: 0.002246113494038582
step: 130, loss: 0.008742082864046097
step: 140, loss: 0.0245478767901659
step: 150, loss: 0.0006595721933990717
step: 160, loss: 0.03332987427711487
step: 170, loss: 0.0008875790517777205
step: 180, loss: 0.006925896275788546
step: 190, loss: 0.0005794812459498644
step: 200, loss: 0.0009309818269684911
step: 210, loss: 0.002559009240940213
/usr/local/lib/python3.7/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=0.44303797468354433, f1=0.510548523206751,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015412485226988792

```



```

step: 10, loss: 0.004476815927773714
step: 20, loss: 0.0017108895117416978
step: 30, loss: 0.0030176681466400623
step: 40, loss: 0.006353017874062061
step: 50, loss: 0.0010536168701946735
step: 60, loss: 0.004732794128358364
step: 70, loss: 0.008356746286153793
step: 80, loss: 0.0009383065043948591
step: 90, loss: 0.040872588753700256
step: 100, loss: 0.0015596768353134394
step: 110, loss: 0.0014418135397136211
step: 120, loss: 0.00663688313215971
step: 130, loss: 0.00030324666295200586
step: 140, loss: 0.0014117341488599777
step: 150, loss: 0.001648225006647408
step: 160, loss: 0.04922918230295181
step: 170, loss: 0.0008393042953684926
step: 180, loss: 0.005056667607277632
step: 190, loss: 0.0010305453324690461
step: 200, loss: 0.001142905093729496
step: 210, loss: 0.0019356774864718318
/usr/local/lib/python3.7/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=0.4353448275862069, f1=0.5052631578947369,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017617283388972282
step: 10, loss: 0.0005784744862467051
step: 20, loss: 0.0007428304525092244
step: 30, loss: 0.00721797626465559
step: 40, loss: 0.0011012105969712138
step: 50, loss: 0.005955886095762253
step: 60, loss: 0.0005946154706180096
step: 70, loss: 0.006273441482335329
step: 80, loss: 0.0011618011631071568
step: 90, loss: 0.02943609096109867
step: 100, loss: 0.0003272672474849969
step: 110, loss: 0.0056977891363203526
step: 120, loss: 0.014605996198952198
step: 130, loss: 0.013655472546815872
step: 140, loss: 0.0009528196533210576
step: 150, loss: 0.00025571201695129275
step: 160, loss: 0.0002573594683781266
step: 170, loss: 0.008182535879313946
step: 180, loss: 0.0021002190187573433

```

```

step: 190, loss: 0.00031685573048889637
step: 200, loss: 0.006232531275600195
step: 210, loss: 0.0005465690628625453
/usr/local/lib/python3.7/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=0.42494226327944573, f1=0.4919540229885057,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006886639748699963
step: 10, loss: 0.0055926209315657616
step: 20, loss: 0.005972490645945072
step: 30, loss: 0.001334338798187673
step: 40, loss: 0.0001623747666599229
step: 50, loss: 0.001482086256146431
step: 60, loss: 0.0003085599164478481
step: 70, loss: 0.0008905419963411987
step: 80, loss: 0.0004037385224364698
step: 90, loss: 0.0003448172938078642
step: 100, loss: 0.0032276292331516743
step: 110, loss: 0.00018947452190332115
step: 120, loss: 0.00030816387152299285
step: 130, loss: 0.002397957956418395
step: 140, loss: 0.04396524280309677
step: 150, loss: 0.00018076406558975577
step: 160, loss: 0.0007345859776251018
step: 170, loss: 0.00011643861944321543
step: 180, loss: 0.0037651644088327885
step: 190, loss: 0.00025704348809085786
step: 200, loss: 0.0015299738151952624
step: 210, loss: 0.002205905271694064
/usr/local/lib/python3.7/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=0.42028985507246375, f1=0.4879227053140097,
best_f1=0.5495652173913044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002809266734402627
step: 10, loss: 0.0003978674940299243
step: 20, loss: 0.03690759465098381
step: 30, loss: 0.014903713949024677
step: 40, loss: 0.0001997423096327111
step: 50, loss: 0.0002579293795861304
step: 60, loss: 0.0056727877818048

```

```

step: 70, loss: 0.001002845005132258
step: 80, loss: 0.0006407427717931569
step: 90, loss: 0.00012081822933396325
step: 100, loss: 0.0005031322361901402
step: 110, loss: 0.0001658390974625945
step: 120, loss: 0.00020013318862766027
step: 130, loss: 0.00011727453238563612
step: 140, loss: 0.00017453249893151224
step: 150, loss: 0.00016063638031482697
step: 160, loss: 0.00025995736359618604
step: 170, loss: 0.002085875952616334
step: 180, loss: 0.000386740401154384
step: 190, loss: 0.026346923783421516
step: 200, loss: 0.006861276458948851
step: 210, loss: 0.0010134493932127953
/usr/local/lib/python3.7/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=0.41964285714285715, f1=0.5022222222222222,
best_f1=0.5495652173913044
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2293it [00:08, 263.21it/s]
load_f1 = 0.5084112149532711
real_f1 = 0.5037593984962406
0it [00:00, ?it/s]/usr/local/lib/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, 177.67it/s]

```

3.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

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

```

Defaults for this optimization level are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5303205251693726
step: 10, loss: 0.3862031400203705
step: 20, loss: 0.3000407814979553
step: 30, loss: 0.4201473891735077
step: 40, loss: 0.4309053421020508
step: 50, loss: 0.30702048540115356
step: 60, loss: 0.30840006470680237
step: 70, loss: 0.29379430413246155
step: 80, loss: 0.23465479910373688
step: 90, loss: 0.30534401535987854
step: 100, loss: 0.2937978208065033
step: 110, loss: 0.5041799545288086
step: 120, loss: 0.11110654473304749
step: 130, loss: 0.186003640294075

```

```

step: 140, loss: 0.06869842112064362
step: 150, loss: 0.1489877700805664
step: 160, loss: 0.06780638545751572
step: 170, loss: 0.1872253268957138
step: 180, loss: 0.02607361227273941
step: 190, loss: 0.35784393548965454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5773195876288659, f1=0.5985037406483791,
best_f1=0.5985037406483791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28591683506965637
step: 10, loss: 0.145869180560112
step: 20, loss: 0.21385493874549866
step: 30, loss: 0.08955886960029602
step: 40, loss: 0.16626335680484772
step: 50, loss: 0.09010424464941025
step: 60, loss: 0.22417403757572174
step: 70, loss: 0.23516686260700226
step: 80, loss: 0.14150165021419525
step: 90, loss: 0.1331881582736969
step: 100, loss: 0.024404115974903107
step: 110, loss: 0.21301217377185822
step: 120, loss: 0.25804221630096436
step: 130, loss: 0.08044274896383286
step: 140, loss: 0.09036274254322052
step: 150, loss: 0.2133331000804901
step: 160, loss: 0.08742763102054596
step: 170, loss: 0.18500211834907532
step: 180, loss: 0.062234655022621155
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.016578562557697296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7300771208226221, f1=0.7411167512690356,
best_f1=0.7411167512690356

```

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

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

```
step: 0, loss: 0.03719903528690338
step: 10, loss: 0.06716660410165787
step: 20, loss: 0.06275662034749985
step: 30, loss: 0.12883122265338898
step: 40, loss: 0.11648838967084885
step: 50, loss: 0.20688562095165253
step: 60, loss: 0.017238186672329903
step: 70, loss: 0.10417954623699188
step: 80, loss: 0.1854383796453476
step: 90, loss: 0.11522508412599564
step: 100, loss: 0.05747559666633606
step: 110, loss: 0.014306631870567799
step: 120, loss: 0.020368224009871483
step: 130, loss: 0.019999828189611435
step: 140, loss: 0.03740323334932327
step: 150, loss: 0.16300293803215027
step: 160, loss: 0.04295025020837784
step: 170, loss: 0.19994889199733734
step: 180, loss: 0.03676706179976463
step: 190, loss: 0.04219254478812218
```

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

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

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

```
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=0.7747252747252747, f1=0.7676240208877285,
best_f1=0.7676240208877285
```

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

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

```
step: 0, loss: 0.05002250894904137
step: 10, loss: 0.17702455818653107
step: 20, loss: 0.03348145633935928
step: 30, loss: 0.09420575946569443
step: 40, loss: 0.19201742112636566
step: 50, loss: 0.09959989786148071
step: 60, loss: 0.08560305088758469
step: 70, loss: 0.15136556327342987
step: 80, loss: 0.08013269305229187
step: 90, loss: 0.004803596995770931
step: 100, loss: 0.2507260739803314
step: 110, loss: 0.00904111098498106
step: 120, loss: 0.07846061885356903
```

```

step: 130, loss: 0.1570775955915451
step: 140, loss: 0.08874472975730896
step: 150, loss: 0.025010602548718452
step: 160, loss: 0.029959874227643013
step: 170, loss: 0.08173482120037079
step: 180, loss: 0.026556646451354027
step: 190, loss: 0.12218278646469116
/usr/local/lib/python3.7/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=0.717557251908397, f1=0.734375, best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07836845517158508
step: 10, loss: 0.0023956485092639923
step: 20, loss: 0.12649273872375488
step: 30, loss: 0.1342843919992447
step: 40, loss: 0.017550788819789886
step: 50, loss: 0.11549561470746994
step: 60, loss: 0.06791265308856964
step: 70, loss: 0.011130064725875854
step: 80, loss: 0.007786758244037628
step: 90, loss: 0.016038991510868073
step: 100, loss: 0.0010965714463964105
step: 110, loss: 0.03381528705358505
step: 120, loss: 0.1822485625743866
step: 130, loss: 0.09884287416934967
step: 140, loss: 0.050422314554452896
step: 150, loss: 0.10322815179824829
step: 160, loss: 0.01159605011343956
step: 170, loss: 0.0012249342398718
step: 180, loss: 0.20252858102321625
step: 190, loss: 0.024417931213974953
/usr/local/lib/python3.7/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=0.7138810198300283, f1=0.7486033519553073,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024625705555081367
step: 10, loss: 0.24287134408950806
step: 20, loss: 0.0025539586786180735
step: 30, loss: 0.01120533887296915
step: 40, loss: 0.05618206784129143
step: 50, loss: 0.04405607655644417

```



```

step: 60, loss: 0.08590944856405258
step: 70, loss: 0.0069051203317940235
step: 80, loss: 0.07901962846517563
step: 90, loss: 0.004332950804382563
step: 100, loss: 0.0003590500564314425
step: 110, loss: 0.0011390276486054063
step: 120, loss: 0.037309419363737106
step: 130, loss: 0.004973005969077349
step: 140, loss: 0.04367877542972565
step: 150, loss: 0.00047690374776721
step: 160, loss: 0.2817081809043884
step: 170, loss: 0.1348668932914734
step: 180, loss: 0.00299365958198905
step: 190, loss: 0.0066113946959376335
/usr/local/lib/python3.7/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=0.7107438016528925, f1=0.7169811320754718,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007772132754325867
step: 10, loss: 0.0021096717100590467
step: 20, loss: 0.020376112312078476
step: 30, loss: 0.012908389791846275
step: 40, loss: 0.003534263465553522
step: 50, loss: 0.11854131519794464
step: 60, loss: 0.007964099757373333
step: 70, loss: 0.00932723842561245
step: 80, loss: 0.0019214511848986149
step: 90, loss: 0.0011272537522017956
step: 100, loss: 0.0005319851916283369
step: 110, loss: 0.008474232628941536
step: 120, loss: 0.0008044381393119693
step: 130, loss: 0.0011215033009648323
step: 140, loss: 0.0030981225427240133
step: 150, loss: 0.017039256170392036
step: 160, loss: 0.01060337945818901
step: 170, loss: 0.000977954943664372
step: 180, loss: 0.006277138367295265
step: 190, loss: 0.0033233186695724726
/usr/local/lib/python3.7/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=0.7211267605633802, f1=0.7450980392156863,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0031475943978875875
step: 10, loss: 0.001936306362040341
step: 20, loss: 0.0009698989451862872
step: 30, loss: 0.0006351750926114619
step: 40, loss: 0.00043408898636698723
step: 50, loss: 0.0003665784897748381
step: 60, loss: 0.0007558532524853945
step: 70, loss: 0.000783981813583523
step: 80, loss: 0.0001367814838886261
step: 90, loss: 0.0008935311925597489
step: 100, loss: 0.0020545090083032846
step: 110, loss: 0.0008867750293575227
step: 120, loss: 0.007057614158838987
step: 130, loss: 0.0006965810898691416
step: 140, loss: 0.0037188134156167507
step: 150, loss: 0.006051268428564072
step: 160, loss: 0.0009397607063874602
step: 170, loss: 0.0006702614482492208
step: 180, loss: 0.003930775914341211
step: 190, loss: 0.00047288535279221833
```

```
/usr/local/lib/python3.7/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=0.7248677248677249, f1=0.7329842931937172,
best_f1=0.7676240208877285
```

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

An input tensor was not cuda.

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

```
step: 0, loss: 0.029140856117010117
step: 10, loss: 0.0005930840270593762
step: 20, loss: 0.024477681145071983
step: 30, loss: 0.002454555593430996
step: 40, loss: 0.0022752743680030107
step: 50, loss: 0.0006252101738937199
step: 60, loss: 0.022429976612329483
step: 70, loss: 0.0003969655663240701
step: 80, loss: 0.0008381074294447899
step: 90, loss: 0.008518459275364876
step: 100, loss: 0.01661757193505764
step: 110, loss: 0.0004394679272081703
step: 120, loss: 0.015053720213472843
step: 130, loss: 0.0005761653883382678
step: 140, loss: 0.0004585389979183674
step: 150, loss: 0.00034478501765988767
step: 160, loss: 0.00019401188183110207
step: 170, loss: 0.00044287709170021117
```

```

step: 180, loss: 0.01349679846316576
step: 190, loss: 0.0004062673542648554
/usr/local/lib/python3.7/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=0.7089947089947092, f1=0.7329842931937172,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014720149338245392
step: 10, loss: 0.0007197451195679605
step: 20, loss: 0.00044452407746575773
step: 30, loss: 0.00012374793004710227
step: 40, loss: 0.0007510874420404434
step: 50, loss: 0.0008586428011767566
step: 60, loss: 0.000316714373184368
step: 70, loss: 0.0009416628745384514
step: 80, loss: 0.00017050221504177898
step: 90, loss: 0.00016449950635433197
step: 100, loss: 0.0003855424583889544
step: 110, loss: 0.00037261380930431187
step: 120, loss: 0.003848637454211712
step: 130, loss: 0.0010659967083483934
step: 140, loss: 0.0002805473341140896
step: 150, loss: 0.0004073463787790388
step: 160, loss: 0.0005921833217144012
step: 170, loss: 0.00020556777599267662
step: 180, loss: 0.0001497876801295206
step: 190, loss: 0.00023716116265859455
/usr/local/lib/python3.7/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=0.7326203208556149, f1=0.7272727272727273,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019972192239947617
step: 10, loss: 0.0003164858790114522
step: 20, loss: 0.0010634275386109948
step: 30, loss: 0.00033548136707395315
step: 40, loss: 0.00029118574457243085
step: 50, loss: 0.0005876439972780645
step: 60, loss: 0.0002188633952755481
step: 70, loss: 0.00020433295867405832
step: 80, loss: 0.0016425995854660869
step: 90, loss: 0.00013918810873292387

```

```

step: 100, loss: 0.0001500679354649037
step: 110, loss: 0.0010789224179461598
step: 120, loss: 0.00029121930128894746
step: 130, loss: 0.0028066642116755247
step: 140, loss: 0.0005200132727622986
step: 150, loss: 0.00012354635691735893
step: 160, loss: 0.00034074895665980875
step: 170, loss: 0.0002529648190829903
step: 180, loss: 0.0002812393067870289
step: 190, loss: 0.0005992516526021063
/usr/local/lib/python3.7/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=0.7302452316076293, f1=0.745945945945946,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019132021407131106
step: 10, loss: 0.0017126970924437046
step: 20, loss: 0.0005218239384703338
step: 30, loss: 0.00033483869628980756
step: 40, loss: 0.000371974369045347
step: 50, loss: 0.0009404337033629417
step: 60, loss: 0.00023959166719578207
step: 70, loss: 0.00018337850633542985
step: 80, loss: 0.0005142401205375791
step: 90, loss: 0.0002977528783958405
step: 100, loss: 0.0010597953805699944
step: 110, loss: 0.00036474273656494915
step: 120, loss: 0.0002583330206107348
step: 130, loss: 0.0008664560155011714
step: 140, loss: 0.0007691021892242134
step: 150, loss: 0.00022881616314407438
step: 160, loss: 9.082483302336186e-05
step: 170, loss: 0.00041959501686505973
step: 180, loss: 0.00012664738460443914
step: 190, loss: 0.00017584397573955357
/usr/local/lib/python3.7/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=0.7422680412371134, f1=0.7538461538461538,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003650809230748564
step: 10, loss: 0.00023627815244253725

```

```

step: 20, loss: 0.002295487094670534
step: 30, loss: 0.0004670721828006208
step: 40, loss: 0.00022722204448655248
step: 50, loss: 0.0005249777459539473
step: 60, loss: 0.0010673115029931068
step: 70, loss: 0.0006123686907812953
step: 80, loss: 0.0009140143520198762
step: 90, loss: 0.0014421935193240643
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.056940071284770966
step: 110, loss: 0.00019708585750777274
step: 120, loss: 0.00015329214511439204
step: 130, loss: 0.00010132935858564451
step: 140, loss: 0.00030247410177253187
step: 150, loss: 9.678552305558696e-05
step: 160, loss: 0.0015889003407210112
step: 170, loss: 0.0002531249774619937
step: 180, loss: 0.00013469357509166002
step: 190, loss: 0.0008530404884368181
/usr/local/lib/python3.7/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=0.7252747252747253, f1=0.7417582417582418,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020041741663590074
step: 10, loss: 0.00012276943016331643
step: 20, loss: 0.00045106065226718783
step: 30, loss: 0.00011553082731552422
step: 40, loss: 0.0002839801018126309
step: 50, loss: 0.000278048450127244
step: 60, loss: 0.00015152375272009522
step: 70, loss: 0.00010983379615936428
step: 80, loss: 0.0009490444208495319
step: 90, loss: 0.0010317665291950107
step: 100, loss: 0.0010051729623228312
step: 110, loss: 0.00023604561283718795
step: 120, loss: 0.0006784358993172646
step: 130, loss: 0.0005158467683941126
step: 140, loss: 0.00018750854360405356
step: 150, loss: 0.00014193844981491566
step: 160, loss: 0.00027085107285529375
step: 170, loss: 0.0003591044805943966
step: 180, loss: 0.00029821021598763764
step: 190, loss: 0.00015628758410457522
/usr/local/lib/python3.7/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=0.7423822714681441, f1=0.7493261455525607,
best_f1=0.7676240208877285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003675827174447477
step: 10, loss: 0.00024071680672932416
step: 20, loss: 0.000159253177116625
step: 30, loss: 0.00048742443323135376
step: 40, loss: 0.00015995411376934499
step: 50, loss: 0.00021423482394311577
step: 60, loss: 0.00015925026673357934
step: 70, loss: 0.00023514157510362566
step: 80, loss: 0.0014494325732812285
step: 90, loss: 9.549382230034098e-05
step: 100, loss: 0.0002686823718249798
step: 110, loss: 8.343060471815988e-05
step: 120, loss: 0.0007505331886932254
step: 130, loss: 0.0008853247272782028
step: 140, loss: 0.0006045598420314491
step: 150, loss: 0.0003964526695199311
step: 160, loss: 0.00024849598412401974
step: 170, loss: 0.00014852384629193693
step: 180, loss: 0.001262521487660706
step: 190, loss: 0.0005585203180089593
/usr/local/lib/python3.7/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=0.7423822714681441, f1=0.745945945945946,
best_f1=0.7676240208877285
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 166.63it/s]
load_f1 = 0.6914893617021276
real_f1 = 0.6494845360824743
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 178.26it/s]

```

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

```

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

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt

```

```
!mv *_dev.pt checkpoints/Abt-Buy.pt

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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



```

UserWarning)
step: 0, loss: 0.6319571733474731
step: 10, loss: 0.3642005920410156
step: 20, loss: 0.2886134088039398
step: 30, loss: 0.3748997151851654
step: 40, loss: 0.27916574478149414
step: 50, loss: 0.26424115896224976
step: 60, loss: 0.3484301269054413
step: 70, loss: 0.36188673973083496
step: 80, loss: 0.31054261326789856
step: 90, loss: 0.17663311958312988
step: 100, loss: 0.24566039443016052
step: 110, loss: 0.21750427782535553
step: 120, loss: 0.07266655564308167
step: 130, loss: 0.057942964136600494
step: 140, loss: 0.20291438698768616
step: 150, loss: 0.22627349197864532
step: 160, loss: 0.12006838619709015
step: 170, loss: 0.27095547318458557
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=0.6370757180156658, f1=0.5861182519280206,
best_f1=0.5861182519280206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25044289231300354
step: 10, loss: 0.298211008310318
step: 20, loss: 0.06463055312633514
step: 30, loss: 0.23173704743385315
step: 40, loss: 0.07754459977149963
step: 50, loss: 0.1530933678150177
step: 60, loss: 0.3556568920612335
step: 70, loss: 0.07889925688505173
step: 80, loss: 0.09676633775234222
step: 90, loss: 0.12659624218940735
step: 100, loss: 0.19526785612106323
step: 110, loss: 0.10268349945545197
step: 120, loss: 0.19877488911151886
step: 130, loss: 0.08647836744785309
step: 140, loss: 0.2153802216053009
step: 150, loss: 0.1808393895626068

```

```

step: 160, loss: 0.14520908892154694
step: 170, loss: 0.04605618864297867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7174447174447175, f1=0.6838407494145199,
best_f1=0.6838407494145199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12001394480466843
step: 10, loss: 0.09213340282440186
step: 20, loss: 0.03474653139710426
step: 30, loss: 0.08562447130680084
step: 40, loss: 0.1216849833726883
step: 50, loss: 0.145889014005661
step: 60, loss: 0.18642063438892365
step: 70, loss: 0.19567571580410004
step: 80, loss: 0.09787274152040482
step: 90, loss: 0.1582944244146347
step: 100, loss: 0.028428778052330017
step: 110, loss: 0.08481057733297348
step: 120, loss: 0.07043958455324173
step: 130, loss: 0.06444214284420013
step: 140, loss: 0.11224721372127533
step: 150, loss: 0.05290435254573822
step: 160, loss: 0.09970194846391678
step: 170, loss: 0.09670715779066086
/usr/local/lib/python3.7/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=0.7154046997389034, f1=0.7146529562982005,
best_f1=0.6838407494145199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01240132562816143
step: 10, loss: 0.13856150209903717
step: 20, loss: 0.05675071105360985
step: 30, loss: 0.13919180631637573
step: 40, loss: 0.008014261722564697
step: 50, loss: 0.18263451755046844
step: 60, loss: 0.19077783823013306
step: 70, loss: 0.061525680124759674

```

```

step: 80, loss: 0.1350145787000656
step: 90, loss: 0.025269145146012306
step: 100, loss: 0.20004849135875702
step: 110, loss: 0.06465837359428406
step: 120, loss: 0.018989797681570053
step: 130, loss: 0.17112784087657928
step: 140, loss: 0.04155602678656578
step: 150, loss: 0.09736144542694092
step: 160, loss: 0.045548539608716965
step: 170, loss: 0.14799964427947998
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7196029776674938, f1=0.6944444444444445,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04172138497233391
step: 10, loss: 0.0347561314702034
step: 20, loss: 0.02965640462934971
step: 30, loss: 0.005770252086222172
step: 40, loss: 0.014603989198803902
step: 50, loss: 0.19737207889556885
step: 60, loss: 0.05465118587017059
step: 70, loss: 0.23210695385932922
step: 80, loss: 0.11997847259044647
step: 90, loss: 0.13537763059139252
step: 100, loss: 0.11445767432451248
step: 110, loss: 0.07785426080226898
step: 120, loss: 0.030599232763051987
step: 130, loss: 0.14433392882347107
step: 140, loss: 0.023885594680905342
step: 150, loss: 0.17060303688049316
step: 160, loss: 0.10382168740034103
step: 170, loss: 0.11144702881574631
/usr/local/lib/python3.7/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=0.6870588235294117, f1=0.7069351230425055,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.011222245171666145
step: 10, loss: 0.08977460116147995
step: 20, loss: 0.008689038455486298
step: 30, loss: 0.02224016934633255
step: 40, loss: 0.06284385919570923
step: 50, loss: 0.1214369609951973
step: 60, loss: 0.04183569177985191
step: 70, loss: 0.03932677209377289
step: 80, loss: 0.045384861528873444
step: 90, loss: 0.009358972311019897
step: 100, loss: 0.03385908901691437
step: 110, loss: 0.002053172793239355
step: 120, loss: 0.007391799241304398
step: 130, loss: 0.007926597259938717
step: 140, loss: 0.19589251279830933
step: 150, loss: 0.051615044474601746
step: 160, loss: 0.062180086970329285
step: 170, loss: 0.026266802102327347
/usr/local/lib/python3.7/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=0.6870229007633587, f1=0.6794258373205743,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00482971454039216
step: 10, loss: 0.03930843994021416
step: 20, loss: 0.004694982897490263
step: 30, loss: 0.03468982130289078
step: 40, loss: 0.06383955478668213
step: 50, loss: 0.05290244147181511
step: 60, loss: 0.00425699632614851
step: 70, loss: 0.007998989894986153
step: 80, loss: 0.018730632960796356
step: 90, loss: 0.0011357904877513647
step: 100, loss: 0.1554650068283081
step: 110, loss: 0.01382453739643097
step: 120, loss: 0.004047485999763012
step: 130, loss: 0.10872554779052734
step: 140, loss: 0.009338350966572762
step: 150, loss: 0.01039922796189785
step: 160, loss: 0.0032371897250413895
step: 170, loss: 0.04632631316781044
/usr/local/lib/python3.7/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=0.6802030456852792, f1=0.6617283950617283,

```

```

best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015588066540658474
step: 10, loss: 0.006026907358318567
step: 20, loss: 0.02728433907032013
step: 30, loss: 0.3408445417881012
step: 40, loss: 0.017038894817233086
step: 50, loss: 0.0027912179939448833
step: 60, loss: 0.023671770468354225
step: 70, loss: 0.036590322852134705
step: 80, loss: 0.0035704155452549458
step: 90, loss: 0.0029918733052909374
step: 100, loss: 0.1708890199661255
step: 110, loss: 0.0826750174164772
step: 120, loss: 0.01623537577688694
step: 130, loss: 0.03542787581682205
step: 140, loss: 0.012791801244020462
step: 150, loss: 0.008371589705348015
step: 160, loss: 0.006677614059299231
step: 170, loss: 0.007266081869602203
/usr/local/lib/python3.7/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=0.6969696969696969, f1=0.7070217917675544,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038514535408467054
step: 10, loss: 0.003993174061179161
step: 20, loss: 0.004283763002604246
step: 30, loss: 0.0756089985370636
step: 40, loss: 0.17921939492225647
step: 50, loss: 0.0010539545910432935
step: 60, loss: 0.03149138763546944
step: 70, loss: 0.0012737520737573504
step: 80, loss: 0.010785059072077274
step: 90, loss: 0.01760842837393284
step: 100, loss: 0.07643269002437592
step: 110, loss: 0.001895107445307076
step: 120, loss: 0.026085250079631805
step: 130, loss: 0.06264110654592514
step: 140, loss: 0.03546600416302681
step: 150, loss: 0.003103244351223111
step: 160, loss: 0.03614344820380211
step: 170, loss: 0.1723269820213318

```

```

/usr/local/lib/python3.7/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=0.707774798927614, f1=0.6909090909090909,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05770310014486313
step: 10, loss: 0.004403893835842609
step: 20, loss: 0.13230766355991364
step: 30, loss: 0.09531198441982269
step: 40, loss: 0.0033555258996784687
step: 50, loss: 0.07760195434093475
step: 60, loss: 0.005199769511818886
step: 70, loss: 0.004897296894341707
step: 80, loss: 0.0035691228695213795
step: 90, loss: 0.007954183034598827
step: 100, loss: 0.006685251370072365
step: 110, loss: 0.019544582813978195
step: 120, loss: 0.0024146686773747206
step: 130, loss: 0.009498139843344688
step: 140, loss: 0.05485466122627258
step: 150, loss: 0.07967973500490189
step: 160, loss: 0.0011161435395479202
step: 170, loss: 0.0008185422630049288
/usr/local/lib/python3.7/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=0.7010309278350515, f1=0.683291770573566,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002428429201245308
step: 10, loss: 0.06697893142700195
step: 20, loss: 0.0012822604039683938
step: 30, loss: 0.001683258917182684
step: 40, loss: 0.00044581954716704786
step: 50, loss: 0.0055776676163077354
step: 60, loss: 0.00582482572644949
step: 70, loss: 0.0004727877676486969
step: 80, loss: 0.001293362583965063
step: 90, loss: 0.0022289101034402847
step: 100, loss: 0.0003740156826097518
step: 110, loss: 0.033918749541044235
step: 120, loss: 0.00044297578278928995
step: 130, loss: 0.00032558635575696826

```

```

step: 140, loss: 0.0025650751776993275
step: 150, loss: 0.0006674132309854031
step: 160, loss: 0.017316052690148354
step: 170, loss: 0.02593991719186306
/usr/local/lib/python3.7/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=0.6868686868686869, f1=0.6714975845410628,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007277446798980236
step: 10, loss: 0.0008259073947556317
step: 20, loss: 0.0037351769860833883
step: 30, loss: 0.00030420266557484865
step: 40, loss: 0.008324508555233479
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.09688466042280197
step: 60, loss: 0.0012775642098858953
step: 70, loss: 0.043593116104602814
step: 80, loss: 0.0003248131542932242
step: 90, loss: 0.0017907473957166076
step: 100, loss: 0.0010514426976442337
step: 110, loss: 0.011487002484500408
step: 120, loss: 0.016364719718694687
step: 130, loss: 0.0018254731548950076
step: 140, loss: 0.009686280973255634
step: 150, loss: 0.001271773362532258
step: 160, loss: 0.0005203665350563824
step: 170, loss: 0.0006945973145775497
/usr/local/lib/python3.7/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=0.6723646723646722, f1=0.6595174262734584,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004469336417969316
step: 10, loss: 0.00238011940382421
step: 20, loss: 0.00912928581237793
step: 30, loss: 0.0001812233531381935
step: 40, loss: 0.0004902396467514336
step: 50, loss: 0.00023555349616799504
step: 60, loss: 0.0015382496640086174
step: 70, loss: 0.004656173288822174
step: 80, loss: 0.0011457190848886967

```

```

step: 90, loss: 0.0001430170377716422
step: 100, loss: 0.2115388661623001
step: 110, loss: 0.00018275990441907197
step: 120, loss: 0.014756505377590656
step: 130, loss: 0.00023184504243545234
step: 140, loss: 0.003160382155328989
step: 150, loss: 0.006324188318103552
step: 160, loss: 0.004628084134310484
step: 170, loss: 0.001195740420371294
/usr/local/lib/python3.7/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=0.6724137931034483, f1=0.6880000000000001,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007289705099537969
step: 10, loss: 0.0008844103431329131
step: 20, loss: 0.008091759867966175
step: 30, loss: 0.07060319930315018
step: 40, loss: 0.0002403081307420507
step: 50, loss: 0.0007927267579361796
step: 60, loss: 0.0003296229406259954
step: 70, loss: 0.00023612195218447596
step: 80, loss: 0.024735160171985626
step: 90, loss: 0.00011388210987206548
step: 100, loss: 0.00036004031426273286
step: 110, loss: 0.00020202971063554287
step: 120, loss: 0.0006544529460370541
step: 130, loss: 0.0003088151279371232
step: 140, loss: 0.001145881717093289
step: 150, loss: 0.0013433977728709579
step: 160, loss: 9.800514089874923e-05
step: 170, loss: 0.0001243258302565664
/usr/local/lib/python3.7/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=0.6721763085399449, f1=0.6910994764397905,
best_f1=0.6944444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002881412801798433
step: 10, loss: 0.0009140034671872854
step: 20, loss: 0.06770621240139008
step: 30, loss: 0.0004738546849694103
step: 40, loss: 0.0011736896121874452

```



```

step: 50, loss: 0.000651294132694602
step: 60, loss: 0.040954262018203735
step: 70, loss: 0.005705642513930798
step: 80, loss: 0.0004797691071871668
step: 90, loss: 0.00021256966283544898
step: 100, loss: 0.006990406196564436
step: 110, loss: 0.0004924663226120174
step: 120, loss: 0.028711970895528793
step: 130, loss: 0.0009108165395446122
step: 140, loss: 0.00020972049969714135
step: 150, loss: 0.01640263944864273
step: 160, loss: 0.12863659858703613
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.0002696315059438348
/usr/local/lib/python3.7/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=0.6703296703296703, f1=0.6822916666666666,
best_f1=0.6944444444444445
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

    warnings.warn("An input tensor was not cuda.")
1916it [00:09, 210.61it/s]
load_f1 = 0.7053140096618358
real_f1 = 0.7241379310344829
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 181.87it/s]

```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

```

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

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

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

```

Selected optimization level O2: FP16 training with FP32 batchnorm and FP32

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

```
step: 0, loss: 0.6031379699707031
step: 10, loss: 0.6395144462585449
step: 20, loss: 0.4851710796356201
step: 30, loss: 0.4518936276435852
step: 40, loss: 0.28168749809265137
step: 50, loss: 0.08428388088941574
step: 60, loss: 0.08016380667686462
step: 70, loss: 0.2860534191131592
step: 80, loss: 0.07184000313282013
step: 90, loss: 0.14025017619132996
step: 100, loss: 0.009048317559063435
step: 110, loss: 0.11347726732492447
step: 120, loss: 0.0385235920548439
```

```

step: 130, loss: 0.015684951096773148
step: 140, loss: 0.02152738906443119
step: 150, loss: 0.06368913501501083
step: 160, loss: 0.04158054292201996
step: 170, loss: 0.24465620517730713
step: 180, loss: 0.08318504691123962
step: 190, loss: 0.01856108382344246
step: 200, loss: 0.09117037802934647
step: 210, loss: 0.013791278935968876
step: 220, loss: 0.02512577921152115
step: 230, loss: 0.19296960532665253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.941834451901566, f1=0.9362186788154897,
best_f1=0.9362186788154897
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01392364501953125
step: 10, loss: 0.03160727396607399
step: 20, loss: 0.2596651017665863
step: 30, loss: 0.15986406803131104
step: 40, loss: 0.13687764108181
step: 50, loss: 0.012287762016057968
step: 60, loss: 0.016190823167562485
step: 70, loss: 0.05321013927459717
step: 80, loss: 0.04578004404902458
step: 90, loss: 0.2633776366710663
step: 100, loss: 0.023373985663056374
step: 110, loss: 0.19603654742240906
step: 120, loss: 0.13089410960674286
step: 130, loss: 0.09958405047655106
step: 140, loss: 0.0037621057126671076
step: 150, loss: 0.05480048805475235
step: 160, loss: 0.014990024268627167
step: 170, loss: 0.007407966535538435
step: 180, loss: 0.04076734185218811
step: 190, loss: 0.0052473219111561775
step: 200, loss: 0.03724247217178345
step: 210, loss: 0.0038774877320975065
step: 220, loss: 0.0744168683886528
step: 230, loss: 0.222891703248024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9552572706935123, f1=0.9410430839002268,
best_f1=0.9410430839002268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028919517993927002
step: 10, loss: 0.022276056930422783
step: 20, loss: 0.12418389320373535
step: 30, loss: 0.012226833961904049
step: 40, loss: 0.04972351714968681
step: 50, loss: 0.10197560489177704
step: 60, loss: 0.004376474767923355
step: 70, loss: 0.04562724009156227
step: 80, loss: 0.0013634926872327924
step: 90, loss: 0.11751794070005417
step: 100, loss: 0.03739327937364578
step: 110, loss: 0.0056849876418709755
step: 120, loss: 0.009360251948237419
step: 130, loss: 0.003514124546200037
step: 140, loss: 0.05623365566134453
step: 150, loss: 0.023737885057926178
step: 160, loss: 0.019434882327914238
step: 170, loss: 0.005188248120248318
step: 180, loss: 0.08724739402532578
step: 190, loss: 0.003070216393098235
step: 200, loss: 0.07335541397333145
step: 210, loss: 0.024481039494276047
step: 220, loss: 0.030829355120658875
step: 230, loss: 0.0015570969553664327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9601769911504424, f1=0.951739618406285,
best_f1=0.951739618406285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039375293999910355

```

```

step: 10, loss: 0.0019487030804157257
step: 20, loss: 0.0023363411892205477
step: 30, loss: 0.0020748923998326063
step: 40, loss: 0.0024054802488535643
step: 50, loss: 0.0005806658882647753
step: 60, loss: 0.0010324738686904311
step: 70, loss: 0.0014774785377085209
step: 80, loss: 0.001110757701098919
step: 90, loss: 0.0017707105726003647
step: 100, loss: 0.00398022448644042
step: 110, loss: 0.0023957001976668835
step: 120, loss: 0.004990576300770044
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.0635017454624176
step: 140, loss: 0.0026253615505993366
step: 150, loss: 0.21686267852783203
step: 160, loss: 0.019668100401759148
step: 170, loss: 0.04588816314935684
step: 180, loss: 0.0006536812870763242
step: 190, loss: 0.0023462011013180017
step: 200, loss: 0.02726718969643116
step: 210, loss: 0.00268554431386292
step: 220, loss: 0.0039055675733834505
step: 230, loss: 0.08697099983692169
/usr/local/lib/python3.7/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=0.9600886917960089, f1=0.9496080627099663,
best_f1=0.951739618406285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019286867463961244
step: 10, loss: 0.0006001087604090571
step: 20, loss: 0.003834584262220516
step: 30, loss: 0.0008806852856650949
step: 40, loss: 0.0017882243264466524
step: 50, loss: 0.0007032065186649561
step: 60, loss: 0.21346797049045563
step: 70, loss: 0.001215155702084303
step: 80, loss: 0.0013435548171401024
step: 90, loss: 0.0031858193688094616
step: 100, loss: 0.0010748454369604588
step: 110, loss: 0.0008938164100982249
step: 120, loss: 0.00027602899353951216
step: 130, loss: 0.007652962114661932
step: 140, loss: 0.005954355001449585
step: 150, loss: 0.0010108217829838395

```

```

step: 160, loss: 0.0010350244119763374
step: 170, loss: 0.007319845724850893
step: 180, loss: 0.050428036600351334
step: 190, loss: 0.13629329204559326
step: 200, loss: 0.005918328650295734
step: 210, loss: 0.002503909170627594
step: 220, loss: 0.015515236184000969
step: 230, loss: 0.0011204632464796305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9603624009060023, f1=0.9450800915331807,
best_f1=0.9450800915331807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004043166060000658
step: 10, loss: 0.007819402031600475
step: 20, loss: 0.005204575601965189
step: 30, loss: 0.0004511185979936272
step: 40, loss: 0.0005584486061707139
step: 50, loss: 0.0009124736534431577
step: 60, loss: 0.0004108596476726234
step: 70, loss: 0.020010508596897125
step: 80, loss: 0.0028338709380477667
step: 90, loss: 0.004965753294527531
step: 100, loss: 0.02672550641000271
step: 110, loss: 0.0013844526838511229
step: 120, loss: 0.0006614943267777562
step: 130, loss: 0.0010884266812354326
step: 140, loss: 0.0002813464670907706
step: 150, loss: 0.00704239634796977
step: 160, loss: 0.0013932830188423395
step: 170, loss: 0.0005820182850584388
step: 180, loss: 0.002916609635576606
step: 190, loss: 0.0003840719582512975
step: 200, loss: 0.0004893068107776344
step: 210, loss: 0.00037068116944283247
step: 220, loss: 0.0005634092376567423
step: 230, loss: 0.09922551363706589
/usr/local/lib/python3.7/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=0.9585666293393057, f1=0.9537767756482525,

```

```

best_f1=0.9450800915331807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003188162110745907
step: 10, loss: 0.0012470112415030599
step: 20, loss: 0.000975678616669029
step: 30, loss: 0.0005006862338632345
step: 40, loss: 0.0013225047150626779
step: 50, loss: 0.000954324146732688
step: 60, loss: 0.0006842829752713442
step: 70, loss: 0.026211746037006378
step: 80, loss: 0.002734798239544034
step: 90, loss: 0.0002244873030576855
step: 100, loss: 0.004522125702351332
step: 110, loss: 0.00033624196657910943
step: 120, loss: 0.0022377350833266973
step: 130, loss: 0.0002950537600554526
step: 140, loss: 0.002040265826508403
step: 150, loss: 0.0012544832425191998
step: 160, loss: 0.1018514335155487
step: 170, loss: 0.023466939106583595
step: 180, loss: 0.10583256930112839
step: 190, loss: 0.0020854880567640066
step: 200, loss: 0.06679078191518784
step: 210, loss: 0.0036171083338558674
step: 220, loss: 0.07512229681015015
step: 230, loss: 0.0022499996703118086
/usr/local/lib/python3.7/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=0.9572072072072072, f1=0.9453302961275627,
best_f1=0.9450800915331807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012627814430743456
step: 10, loss: 0.000859764579217881
step: 20, loss: 0.0004961717058904469
step: 30, loss: 0.0006523014162667096
step: 40, loss: 0.005670711863785982
step: 50, loss: 0.00042413995834067464
step: 60, loss: 0.0008143538143485785
step: 70, loss: 0.0010022082133218646
step: 80, loss: 0.0035793909337371588
step: 90, loss: 0.00028631248278543353
step: 100, loss: 0.0017403494566679
step: 110, loss: 0.004537269007414579

```



```

step: 120, loss: 0.00258272304199636
step: 130, loss: 0.0023533774074167013
step: 140, loss: 0.00039865588769316673
step: 150, loss: 0.00020083699200768024
step: 160, loss: 0.0003131064586341381
step: 170, loss: 0.000551500532310456
step: 180, loss: 0.0003776292141992599
step: 190, loss: 0.002708052285015583
step: 200, loss: 0.00044406281085684896
step: 210, loss: 0.0003240909136366099
step: 220, loss: 0.00039580839802511036
step: 230, loss: 0.00019320538558531553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9615384615384616, f1=0.9437428243398392,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023468512517865747
step: 10, loss: 0.0006079564918763936
step: 20, loss: 0.000907513196580112
step: 30, loss: 6.208675767993554e-05
step: 40, loss: 0.056735455989837646
step: 50, loss: 7.380809256574139e-05
step: 60, loss: 0.00032711762469261885
step: 70, loss: 0.0009654348250478506
step: 80, loss: 0.00016059599875006825
step: 90, loss: 0.0034859429579228163
step: 100, loss: 0.00025212904438376427
step: 110, loss: 0.00018908927449956536
step: 120, loss: 0.0004856935702264309
step: 130, loss: 0.00025399221340194345
step: 140, loss: 0.00023822631919756532
step: 150, loss: 0.012500963173806667
step: 160, loss: 0.0008735501323826611
step: 170, loss: 0.008813789114356041
step: 180, loss: 0.0021541116293519735
step: 190, loss: 0.0008190224762074649
step: 200, loss: 0.0009085024357773364
step: 210, loss: 0.0009267203859053552
step: 220, loss: 0.0006455217371694744
step: 230, loss: 0.00804334506392479

```

```

/usr/local/lib/python3.7/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=0.9578713968957872, f1=0.9454949944382648,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003699498192872852
step: 10, loss: 0.0014034447958692908
step: 20, loss: 0.0002205623168265447
step: 30, loss: 0.00028324511367827654
step: 40, loss: 0.00019363056344445795
step: 50, loss: 0.02029695361852646
step: 60, loss: 0.000189742335351184
step: 70, loss: 0.0002371486771153286
step: 80, loss: 0.002441721735522151
step: 90, loss: 0.0057478975504636765
step: 100, loss: 0.00020186665642540902
step: 110, loss: 0.00021525412739720196
step: 120, loss: 0.0002586777263786644
step: 130, loss: 0.0004943434614688158
step: 140, loss: 0.04362647980451584
step: 150, loss: 0.016739021986722946
step: 160, loss: 0.00014007158461026847
step: 170, loss: 7.493374141631648e-05
step: 180, loss: 0.00033846095902845263
step: 190, loss: 0.0008006296120584011
step: 200, loss: 0.0003533612471073866
step: 210, loss: 4.122514292248525e-05
step: 220, loss: 0.00015730602899566293
step: 230, loss: 0.0001514967152616009
/usr/local/lib/python3.7/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=0.9611542730299667, f1=0.9513274336283186,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0062250131741166115
step: 10, loss: 0.0006065224879421294
step: 20, loss: 0.00017329270485788584
step: 30, loss: 0.00792425312101841
step: 40, loss: 0.00023545343719888479
step: 50, loss: 0.00023073684133123606
step: 60, loss: 0.0034478437155485153
step: 70, loss: 0.0022673809435218573

```

```

step: 80, loss: 9.845834574662149e-05
step: 90, loss: 0.0001150895914179273
step: 100, loss: 0.00039914369699545205
step: 110, loss: 0.00031606011907570064
step: 120, loss: 0.0006419881829060614
step: 130, loss: 0.0012686800910159945
step: 140, loss: 0.0005549454363062978
step: 150, loss: 0.0017262604087591171
step: 160, loss: 0.0001311220257775858
step: 170, loss: 0.0006354896468110383
step: 180, loss: 0.0003704876289702952
step: 190, loss: 0.00043940701289102435
step: 200, loss: 0.00024558737641200423
step: 210, loss: 7.792279939167202e-05
step: 220, loss: 0.00019943166989833117
step: 230, loss: 0.0002039407117990777
/usr/local/lib/python3.7/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=0.9543937708565072, f1=0.9452513966480447,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025177127099595964
step: 10, loss: 7.98700493760407e-05
step: 20, loss: 0.007781794294714928
step: 30, loss: 0.006342113018035889
step: 40, loss: 0.00023988046450540423
step: 50, loss: 0.000277009850833565
step: 60, loss: 0.00023004625109024346
step: 70, loss: 0.00010934981401078403
step: 80, loss: 0.00016258454706985503
step: 90, loss: 0.00014099245890974998
step: 100, loss: 6.824478623457253e-05
step: 110, loss: 0.00010182779078604653
step: 120, loss: 0.00015427192556671798
step: 130, loss: 0.0002072320057777688
step: 140, loss: 0.0001259842683793977
step: 150, loss: 0.00010593779006740078
step: 160, loss: 0.00017398083582520485
step: 170, loss: 0.0001303239696426317
step: 180, loss: 0.010336140170693398
step: 190, loss: 9.249307186109945e-05
step: 200, loss: 0.00014221682795323431
step: 210, loss: 0.0002304171648574993
step: 220, loss: 0.00018564797937870026
step: 230, loss: 0.028775779530405998

```

```

/usr/local/lib/python3.7/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=0.9567147613762487, f1=0.9519553072625698,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004086430184543133
step: 10, loss: 0.0007963841198943555
step: 20, loss: 0.0008389002177864313
step: 30, loss: 0.00012562074698507786
step: 40, loss: 0.0002347846020711586
step: 50, loss: 0.00015952711692079902
step: 60, loss: 0.0008259573951363564
step: 70, loss: 7.212725904537365e-05
step: 80, loss: 0.00010146988643100485
step: 90, loss: 0.0001696472172625363
step: 100, loss: 0.00010635569924488664
step: 110, loss: 0.0035049046855419874
step: 120, loss: 0.0002281602646689862
step: 130, loss: 0.00015510397497564554
step: 140, loss: 0.00016193880583159626
step: 150, loss: 0.0010105184046551585
step: 160, loss: 0.00034558004699647427
step: 170, loss: 0.000133568377350457
step: 180, loss: 9.025300096254796e-05
step: 190, loss: 0.00017394727910868824
step: 200, loss: 8.380268991459161e-05
step: 210, loss: 4.033757795696147e-05
step: 220, loss: 0.00021690120047423989
step: 230, loss: 7.22959084669128e-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=0.9553571428571428, f1=0.946067415730337,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002517151879146695
step: 10, loss: 0.0001310377847403288
step: 20, loss: 0.0006152056157588959
step: 30, loss: 0.00012680416693910956
step: 40, loss: 0.0026849096175283194
step: 50, loss: 9.884955215966329e-05
step: 60, loss: 0.00011550471390364692
step: 70, loss: 0.005377591587603092

```

```

step: 80, loss: 0.0002115554379997775
step: 90, loss: 0.0002508146280888468
step: 100, loss: 7.46311925468035e-05
step: 110, loss: 7.038266630843282e-05
step: 120, loss: 4.307348353904672e-05
step: 130, loss: 0.001676828251220286
step: 140, loss: 8.454519411316141e-05
step: 150, loss: 5.854406117578037e-05
step: 160, loss: 0.0016529355198144913
step: 170, loss: 0.00017345805827062577
step: 180, loss: 0.00019177993817720562
step: 190, loss: 7.589710730826482e-05
step: 200, loss: 5.106938260723837e-05
step: 210, loss: 6.937274883966893e-05
step: 220, loss: 6.615062739001587e-05
step: 230, loss: 0.00490990374237299
/usr/local/lib/python3.7/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=0.9599109131403119, f1=0.9499443826473859,
best_f1=0.9437428243398392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.428950731176883e-05
step: 10, loss: 6.059036240912974e-05
step: 20, loss: 0.0003290711611043662
step: 30, loss: 0.00010443662176840007
step: 40, loss: 0.00019685234292410314
step: 50, loss: 0.04553942382335663
step: 60, loss: 0.00023243596660904586
step: 70, loss: 0.0010984324617311358
step: 80, loss: 8.461552351946011e-05
step: 90, loss: 8.522660937160254e-05
step: 100, loss: 6.954609125386924e-05
step: 110, loss: 9.543069609208032e-05
step: 120, loss: 9.890382352750748e-05
step: 130, loss: 7.656248635612428e-05
step: 140, loss: 4.846606680075638e-05
step: 150, loss: 0.0002769990242086351
step: 160, loss: 4.3631080188788474e-05
step: 170, loss: 7.443955837516114e-05
step: 180, loss: 4.1917654016288e-05
step: 190, loss: 0.00031894363928586245
step: 200, loss: 0.0003086219367105514
step: 210, loss: 0.0004757953865919262
step: 220, loss: 0.00012162249186076224
step: 230, loss: 7.032047869870439e-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=0.9588431590656283, f1=0.9500554938956715,
best_f1=0.9437428243398392
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 192.12it/s]
load_f1 = 0.9620535714285715
real_f1 = 0.9577777777777777
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 235.00it/s]

```

3.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale              : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6345385313034058
step: 10, loss: 0.5159788727760315
step: 20, loss: 0.5931004285812378
step: 30, loss: 0.3431259095668793
step: 40, loss: 0.3452569842338562
step: 50, loss: 0.2996799647808075
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: 60, loss: 0.22765102982521057
step: 70, loss: 0.23037375509738922
step: 80, loss: 0.1413046270608902
step: 90, loss: 0.31314384937286377
step: 100, loss: 0.22369693219661713
step: 110, loss: 0.09800390154123306
step: 120, loss: 0.13428767025470734
step: 130, loss: 0.1631496399641037
step: 140, loss: 0.0863686129450798
step: 150, loss: 0.05076254904270172
step: 160, loss: 0.0705532357096672
step: 170, loss: 0.2859067916870117
step: 180, loss: 0.05679988116025925
step: 190, loss: 0.019781971350312233
step: 200, loss: 0.17236925661563873
step: 210, loss: 0.0519067719578743
step: 220, loss: 0.20721620321273804

```



```

step: 230, loss: 0.13942179083824158
step: 240, loss: 0.09976700693368912
step: 250, loss: 0.09312956780195236
step: 260, loss: 0.062252700328826904
step: 270, loss: 0.04090512543916702
step: 280, loss: 0.18595346808433533
step: 290, loss: 0.30160561203956604
step: 300, loss: 0.10853982716798782
step: 310, loss: 0.18380101025104523
step: 320, loss: 0.07394029945135117
step: 330, loss: 0.04933370277285576
step: 340, loss: 0.17994646728038788
step: 350, loss: 0.06874911487102509
step: 360, loss: 0.09396405518054962
step: 370, loss: 0.10824308544397354
step: 380, loss: 0.0333951972424984
step: 390, loss: 0.14608217775821686
step: 400, loss: 0.2794446349143982
step: 410, loss: 0.0994172990322113
step: 420, loss: 0.16915026307106018
step: 430, loss: 0.20114783942699432
step: 440, loss: 0.09981073439121246
step: 450, loss: 0.00991132203489542
step: 460, loss: 0.08980310708284378
step: 470, loss: 0.11010853201150894
step: 480, loss: 0.14469105005264282
step: 490, loss: 0.2282792627811432
step: 500, loss: 0.09226910024881363
step: 510, loss: 0.14934997260570526
step: 520, loss: 0.2697022557258606
step: 530, loss: 0.02729497291147709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9067641681901281, f1=0.8931860036832413,
best_f1=0.8931860036832413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1297970712184906
step: 10, loss: 0.12176104635000229
step: 20, loss: 0.10667318105697632
step: 30, loss: 0.1161021739244461
step: 40, loss: 0.04303897172212601

```

step: 50, loss: 0.07592345774173737
step: 60, loss: 0.049467604607343674
step: 70, loss: 0.05676442012190819
step: 80, loss: 0.06961308419704437
step: 90, loss: 0.05212937667965889
step: 100, loss: 0.05721186846494675
step: 110, loss: 0.024939224123954773
step: 120, loss: 0.05967044457793236
step: 130, loss: 0.10777202248573303
step: 140, loss: 0.10405401885509491
step: 150, loss: 0.05712760239839554
step: 160, loss: 0.03692019730806351
step: 170, loss: 0.09184528887271881
step: 180, loss: 0.010978949256241322
step: 190, loss: 0.05827793478965759
step: 200, loss: 0.024434037506580353
step: 210, loss: 0.09479187428951263
step: 220, loss: 0.04099050164222717
step: 230, loss: 0.04752489551901817
step: 240, loss: 0.028789175674319267
step: 250, loss: 0.22742146253585815
step: 260, loss: 0.024253079667687416
step: 270, loss: 0.24112984538078308
step: 280, loss: 0.1962643712759018
step: 290, loss: 0.028782367706298828
step: 300, loss: 0.0925590842962265
step: 310, loss: 0.03146238252520561
step: 320, loss: 0.1843540519475937
step: 330, loss: 0.19272997975349426
step: 340, loss: 0.07221384346485138
step: 350, loss: 0.03916364535689354
step: 360, loss: 0.09560562670230865
step: 370, loss: 0.19742199778556824
step: 380, loss: 0.11260980367660522
step: 390, loss: 0.19608673453330994
step: 400, loss: 0.02135448344051838
step: 410, loss: 0.09618993103504181
step: 420, loss: 0.026998300105333328
step: 430, loss: 0.22829782962799072
step: 440, loss: 0.12894001603126526
step: 450, loss: 0.05193041265010834
step: 460, loss: 0.09953662008047104
step: 470, loss: 0.03979341685771942
step: 480, loss: 0.15147246420383453
step: 490, loss: 0.03275294229388237
step: 500, loss: 0.5400466322898865
step: 510, loss: 0.0367853119969368
step: 520, loss: 0.12258786708116531

```

step: 530, loss: 0.09555791318416595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.908235294117647, f1=0.9002364066193854,
best_f1=0.9002364066193854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0767800360918045
step: 10, loss: 0.09259124845266342
step: 20, loss: 0.0495271235704422
step: 30, loss: 0.26746612787246704
step: 40, loss: 0.0230867899954319
step: 50, loss: 0.16216401755809784
step: 60, loss: 0.05473768338561058
step: 70, loss: 0.03925517946481705
step: 80, loss: 0.006715036928653717
step: 90, loss: 0.04146776348352432
step: 100, loss: 0.02179226279258728
step: 110, loss: 0.002379094250500202
step: 120, loss: 0.01759437844157219
step: 130, loss: 0.004072373267263174
step: 140, loss: 0.017164321616292
step: 150, loss: 0.07640637457370758
step: 160, loss: 0.009216570295393467
step: 170, loss: 0.09696120768785477
step: 180, loss: 0.06885220110416412
step: 190, loss: 0.04847704991698265
step: 200, loss: 0.0153276976197958
step: 210, loss: 0.028813179582357407
step: 220, loss: 0.06390254199504852
step: 230, loss: 0.28438472747802734
step: 240, loss: 0.0031115247402340174
step: 250, loss: 0.03987603634595871
step: 260, loss: 0.0205225870013237
step: 270, loss: 0.022196004167199135
step: 280, loss: 0.2078208327293396
step: 290, loss: 0.007804475259035826
step: 300, loss: 0.12049132585525513
step: 310, loss: 0.035007674247026443
step: 320, loss: 0.042947255074977875
step: 330, loss: 0.015897002071142197
step: 340, loss: 0.010643357411026955

```

```

step: 350, loss: 0.019028298556804657
step: 360, loss: 0.03255254775285721
step: 370, loss: 0.013550067320466042
step: 380, loss: 0.005147553980350494
step: 390, loss: 0.024470722302794456
step: 400, loss: 0.1590900868177414
step: 410, loss: 0.015072374604642391
step: 420, loss: 0.2630383372306824
step: 430, loss: 0.010259999893605709
step: 440, loss: 0.07859810441732407
step: 450, loss: 0.12051332741975784
step: 460, loss: 0.08320201188325882
step: 470, loss: 0.034559376537799835
step: 480, loss: 0.02945765294134617
step: 490, loss: 0.022369258105754852
step: 500, loss: 0.12409017980098724
step: 510, loss: 0.014923984184861183
step: 520, loss: 0.05362401902675629
step: 530, loss: 0.0747075006365776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.912082745651152, f1=0.8944099378881988,
best_f1=0.8944099378881988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008499042131006718
step: 10, loss: 0.008325438015162945
step: 20, loss: 0.017380500212311745
step: 30, loss: 0.009126865305006504
step: 40, loss: 0.04010215401649475
step: 50, loss: 0.03553945943713188
step: 60, loss: 0.006586793344467878
step: 70, loss: 0.004099797457456589
step: 80, loss: 0.003751944750547409
step: 90, loss: 0.1672637015581131
step: 100, loss: 0.00903908722102642
step: 110, loss: 0.15919935703277588
step: 120, loss: 0.0019498032052069902
step: 130, loss: 0.002475582994520664
step: 140, loss: 0.004589851479977369
step: 150, loss: 0.018469564616680145
step: 160, loss: 0.06064090505242348

```

```

step: 170, loss: 0.004836102947592735
step: 180, loss: 0.0012690932489931583
step: 190, loss: 0.03329590708017349
step: 200, loss: 0.009394606575369835
step: 210, loss: 0.046146899461746216
step: 220, loss: 0.0071780625730752945
step: 230, loss: 0.01597127877175808
step: 240, loss: 0.004680955782532692
step: 250, loss: 0.05700678750872612
step: 260, loss: 0.0022197160869836807
step: 270, loss: 0.04535209387540817
step: 280, loss: 0.21195867657661438
step: 290, loss: 0.025626741349697113
step: 300, loss: 0.004644455853849649
step: 310, loss: 0.004963350016623735
step: 320, loss: 0.0108055854216218
step: 330, loss: 0.021290190517902374
step: 340, loss: 0.05742461606860161
step: 350, loss: 0.12710030376911163
step: 360, loss: 0.0900988057255745
step: 370, loss: 0.005927878897637129
step: 380, loss: 0.010796040296554565
step: 390, loss: 0.005154358223080635
step: 400, loss: 0.02862093411386013
step: 410, loss: 0.0008702185587026179
step: 420, loss: 0.08580722659826279
step: 430, loss: 0.06089651584625244
step: 440, loss: 0.0022630123421549797
step: 450, loss: 0.008185058832168579
step: 460, loss: 0.008338574320077896
step: 470, loss: 0.046225808560848236
step: 480, loss: 0.04310920089483261
step: 490, loss: 0.010161244310438633
step: 500, loss: 0.0024882443249225616
step: 510, loss: 0.028400063514709473
step: 520, loss: 0.03319086134433746
step: 530, loss: 0.041965920478105545
/usr/local/lib/python3.7/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=0.9072261072261073, f1=0.8899430740037951,
best_f1=0.8944099378881988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03445682302117348
step: 10, loss: 0.01999519020318985
step: 20, loss: 0.003906242549419403

```

step: 30, loss: 0.004001666326075792
step: 40, loss: 0.04595369100570679
step: 50, loss: 0.024317516013979912
step: 60, loss: 0.09414973855018616
step: 70, loss: 0.024932237342000008
step: 80, loss: 0.008136275224387646
step: 90, loss: 0.02557588554918766
step: 100, loss: 0.0347147136926651
step: 110, loss: 0.0016749163623899221
step: 120, loss: 0.05267760157585144
step: 130, loss: 0.0020467801950871944
step: 140, loss: 0.003363917348906398
step: 150, loss: 0.10932005196809769
step: 160, loss: 0.0008929428295232356
step: 170, loss: 0.031720083206892014
step: 180, loss: 0.0009285951964557171
step: 190, loss: 0.00449293153360486
step: 200, loss: 0.017756709828972816
step: 210, loss: 0.06895583868026733
step: 220, loss: 0.0031013451516628265
step: 230, loss: 0.02513963170349598
step: 240, loss: 0.043276891112327576
step: 250, loss: 0.04618023708462715
step: 260, loss: 0.0025880122557282448
step: 270, loss: 0.010830337181687355
step: 280, loss: 0.007806032430380583
step: 290, loss: 0.208930104970932
step: 300, loss: 0.03937363997101784
step: 310, loss: 0.00972413457930088
step: 320, loss: 0.07521802932024002
step: 330, loss: 0.01639087311923504
step: 340, loss: 0.0031206447165459394
step: 350, loss: 0.0013627088628709316
step: 360, loss: 0.0039001612458378077
step: 370, loss: 0.0029813102446496487
step: 380, loss: 0.004460127558559179
step: 390, loss: 0.00213326932862401
step: 400, loss: 0.0011027674190700054
step: 410, loss: 0.007287752814590931
step: 420, loss: 0.0006139019387774169
step: 430, loss: 0.1486971229314804
step: 440, loss: 0.179969921708107
step: 450, loss: 0.011231125332415104
step: 460, loss: 0.004497077316045761
step: 470, loss: 0.03633009269833565
step: 480, loss: 0.016462404280900955
step: 490, loss: 0.003463469911366701
step: 500, loss: 0.04566683992743492

```

step: 510, loss: 0.008569885976612568
step: 520, loss: 0.00778389535844326
step: 530, loss: 0.0008788554114289582
/usr/local/lib/python3.7/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=0.9111111111111111, f1=0.8932400932400931,
best_f1=0.8944099378881988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004449534695595503
step: 10, loss: 0.0004503014497458935
step: 20, loss: 0.13122524321079254
step: 30, loss: 0.002606580499559641
step: 40, loss: 0.006317175459116697
step: 50, loss: 0.010005541145801544
step: 60, loss: 0.0012756913201883435
step: 70, loss: 0.0009119705064222217
step: 80, loss: 0.014675899408757687
step: 90, loss: 0.0003623013326432556
step: 100, loss: 0.14486517012119293
step: 110, loss: 0.027819698676466942
step: 120, loss: 0.005792984273284674
step: 130, loss: 0.004270556848496199
step: 140, loss: 0.00865299440920353
step: 150, loss: 0.0012288432335481048
step: 160, loss: 0.0019091577269136906
step: 170, loss: 0.001449735020287335
step: 180, loss: 0.0006759064854122698
step: 190, loss: 0.013389325700700283
step: 200, loss: 0.010476130992174149
step: 210, loss: 0.0023250060621649027
step: 220, loss: 0.0010475533781573176
step: 230, loss: 0.0016632708720862865
step: 240, loss: 0.04200168326497078
step: 250, loss: 0.004966967273503542
step: 260, loss: 0.0003428615164011717
step: 270, loss: 0.0013601050013676286
step: 280, loss: 0.0009884071769192815
step: 290, loss: 0.00039075847598724067
step: 300, loss: 0.0004962278180755675
step: 310, loss: 0.0010300506837666035
step: 320, loss: 0.0015178339090198278
step: 330, loss: 0.0014609888894483447
step: 340, loss: 0.0443820022046566
step: 350, loss: 0.0028169522993266582
step: 360, loss: 0.004646338988095522

```

```
step: 370, loss: 0.0031775704119354486
step: 380, loss: 0.000339863559929654
step: 390, loss: 0.004031868651509285
step: 400, loss: 0.14433667063713074
step: 410, loss: 0.0880168229341507
step: 420, loss: 0.02476143091917038
step: 430, loss: 0.0013658857205882668
step: 440, loss: 0.0006460821023210883
step: 450, loss: 0.0005717584281228483
step: 460, loss: 0.0011738475877791643
step: 470, loss: 0.0005673015839420259
step: 480, loss: 0.0011858321959152818
step: 490, loss: 0.007445620372891426
step: 500, loss: 0.0023434069007635117
step: 510, loss: 0.02280420809984207
step: 520, loss: 0.0011516499798744917
step: 530, loss: 0.002705085091292858
/usr/local/lib/python3.7/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=0.9083969465648856, f1=0.8818137964302943,
best_f1=0.8944099378881988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025044314679689705
step: 10, loss: 0.0004061848158016801
step: 20, loss: 0.001335824839770794
step: 30, loss: 0.0008527265163138509
step: 40, loss: 0.00617134477943182
step: 50, loss: 0.0050085559487342834
step: 60, loss: 0.0007466660463251173
step: 70, loss: 0.00033340559457428753
step: 80, loss: 0.005214403383433819
step: 90, loss: 0.0005238659796305001
step: 100, loss: 0.00040051789255812764
step: 110, loss: 0.006995262578129768
step: 120, loss: 0.002682599239051342
step: 130, loss: 0.08341220766305923
step: 140, loss: 6.665248656645417e-05
step: 150, loss: 0.0011732700513675809
step: 160, loss: 0.0001616103108972311
step: 170, loss: 0.0015709124272689223
step: 180, loss: 0.0002570578071754426
step: 190, loss: 0.009640974923968315
step: 200, loss: 0.00040595553582534194
step: 210, loss: 0.014825894497334957
step: 220, loss: 0.00165260280482471
```



```

step: 230, loss: 0.014640148729085922
step: 240, loss: 0.0020671836100518703
step: 250, loss: 0.08286258578300476
step: 260, loss: 0.0002570589422248304
step: 270, loss: 0.0026491787284612656
step: 280, loss: 0.0006049229414202273
step: 290, loss: 0.0009963074699044228
step: 300, loss: 0.2320452481508255
step: 310, loss: 0.0025579421781003475
step: 320, loss: 0.000618698017206043
step: 330, loss: 0.005303935147821903
step: 340, loss: 0.005940405651926994
step: 350, loss: 0.000491646584123373
step: 360, loss: 0.008164425380527973
step: 370, loss: 0.0006243972457014024
step: 380, loss: 0.0002820958034135401
step: 390, loss: 0.00021777309302706271
step: 400, loss: 0.00036427087616175413
step: 410, loss: 0.003521945793181658
step: 420, loss: 0.003456655889749527
step: 430, loss: 0.006967911496758461
step: 440, loss: 0.0007361914031207561
step: 450, loss: 0.019948331639170647
step: 460, loss: 0.006426415406167507
step: 470, loss: 0.00027242579380981624
step: 480, loss: 0.05668797716498375
step: 490, loss: 0.012163490056991577
step: 500, loss: 0.0008651715470477939
step: 510, loss: 0.0012223549420014024
step: 520, loss: 0.06211160123348236
step: 530, loss: 0.0012928241631016135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.913144449605202, f1=0.9038912330051571,
best_f1=0.9038912330051571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002349399495869875
step: 10, loss: 0.0023497575893998146
step: 20, loss: 0.00016778033750597388
step: 30, loss: 0.014491303823888302
step: 40, loss: 0.004725146573036909

```

step: 50, loss: 0.024494972079992294
step: 60, loss: 0.00018562246987130493
step: 70, loss: 0.00013842048065271229
step: 80, loss: 0.008216220885515213
step: 90, loss: 0.00029043335234746337
step: 100, loss: 0.00013454785221256316
step: 110, loss: 0.16483156383037567
step: 120, loss: 0.0006506711943075061
step: 130, loss: 0.00016769388457760215
step: 140, loss: 0.000950803398154676
step: 150, loss: 0.00024873760412447155
step: 160, loss: 0.0022500744089484215
step: 170, loss: 0.0011380473151803017
step: 180, loss: 8.360248466487974e-05
step: 190, loss: 0.002226270502433181
step: 200, loss: 0.0014610171783715487
step: 210, loss: 0.00029705456108786166
step: 220, loss: 0.0014187708729878068
step: 230, loss: 0.0003562140045687556
step: 240, loss: 0.0008180763688869774
step: 250, loss: 0.005781116429716349
step: 260, loss: 0.00038683952880091965
step: 270, loss: 0.005035795271396637
step: 280, loss: 0.027199214324355125
step: 290, loss: 0.03314600884914398
step: 300, loss: 0.0031112851575016975
step: 310, loss: 0.00024271439178846776
step: 320, loss: 0.1319313943386078
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 330, loss: 0.0004132980538997799
step: 340, loss: 0.0009593864087946713
step: 350, loss: 0.0002708043612074107
step: 360, loss: 0.0033416152000427246
step: 370, loss: 0.0012456387048587203
step: 380, loss: 0.005161578301340342
step: 390, loss: 0.017933432012796402
step: 400, loss: 0.0005495662335306406
step: 410, loss: 0.00045497005339711905
step: 420, loss: 0.002170708728954196
step: 430, loss: 0.02802407369017601
step: 440, loss: 0.0022748075425624847
step: 450, loss: 0.006901554763317108
step: 460, loss: 0.00014681287575513124
step: 470, loss: 0.0007365681231021881
step: 480, loss: 0.0002736441674642265
step: 490, loss: 0.0011341498466208577
step: 500, loss: 0.06769923120737076
step: 510, loss: 0.07875583320856094

```

step: 520, loss: 0.003987311851233244
step: 530, loss: 0.00015358885866589844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9187411930483794, f1=0.889947594092425,
best_f1=0.889947594092425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015389852924272418
step: 10, loss: 9.928735380526632e-05
step: 20, loss: 0.0006311328615993261
step: 30, loss: 0.0034648380242288113
step: 40, loss: 0.00025116270990110934
step: 50, loss: 0.00015247953706420958
step: 60, loss: 0.00017920020036399364
step: 70, loss: 4.028594048577361e-05
step: 80, loss: 0.00021500600269064307
step: 90, loss: 0.0013528912095353007
step: 100, loss: 6.209599087014794e-05
step: 110, loss: 4.9695772759150714e-05
step: 120, loss: 0.008853926323354244
step: 130, loss: 0.007361988537013531
step: 140, loss: 0.00017574429512023926
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.0002517862303648144
step: 160, loss: 0.00011120025010313839
step: 170, loss: 0.002388553461059928
step: 180, loss: 0.00021892438235227019
step: 190, loss: 0.002936912002041936
step: 200, loss: 0.013268006965517998
step: 210, loss: 0.009460082277655602
step: 220, loss: 0.0004288560012355447
step: 230, loss: 0.00744528928771615
step: 240, loss: 0.00011219615407753736
step: 250, loss: 0.0004921869840472937
step: 260, loss: 0.0008639739826321602
step: 270, loss: 0.0001238436670973897
step: 280, loss: 0.0028611000161617994
step: 290, loss: 0.009457272477447987
step: 300, loss: 0.00014398052007891238
step: 310, loss: 0.00032785063376650214
step: 320, loss: 0.0002851183235179633

```

```
step: 330, loss: 0.00039976471452973783
step: 340, loss: 0.0043961466290056705
step: 350, loss: 0.07424214482307434
step: 360, loss: 0.00017513500642962754
step: 370, loss: 0.00034524311195127666
step: 380, loss: 0.00020145170856267214
step: 390, loss: 0.003722775261849165
step: 400, loss: 0.0014471783069893718
step: 410, loss: 0.00016248297470156103
step: 420, loss: 9.161894558928907e-05
step: 430, loss: 8.408936992054805e-05
step: 440, loss: 0.00015127783990465105
step: 450, loss: 0.002677884651347995
step: 460, loss: 0.0035219318233430386
step: 470, loss: 0.23010286688804626
step: 480, loss: 0.00015136078582145274
step: 490, loss: 0.0001337494031758979
step: 500, loss: 0.004487516358494759
step: 510, loss: 0.029645023867487907
step: 520, loss: 0.0006073230761103332
step: 530, loss: 0.00010697665129555389
/usr/local/lib/python3.7/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=0.9113805970149255, f1=0.8901408450704226,
best_f1=0.889947594092425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021461817959789187
step: 10, loss: 4.9917580327019095e-05
step: 20, loss: 0.00012318379594944417
step: 30, loss: 0.004075329750776291
step: 40, loss: 0.00019366535707376897
step: 50, loss: 0.00015605389489792287
step: 60, loss: 9.962977492250502e-05
step: 70, loss: 7.408553938148543e-05
step: 80, loss: 0.00013824008055962622
step: 90, loss: 0.001672322046943009
step: 100, loss: 8.287890523206443e-05
step: 110, loss: 5.671848703059368e-05
step: 120, loss: 0.0012600445188581944
step: 130, loss: 0.00018152942357119173
step: 140, loss: 0.00013959530042484403
step: 150, loss: 0.0001847779203671962
step: 160, loss: 8.261114271590486e-05
step: 170, loss: 0.00039136307896114886
step: 180, loss: 0.00022743834415450692
```

```
step: 190, loss: 0.0005675561842508614
step: 200, loss: 0.0007492689765058458
step: 210, loss: 0.0008244640193879604
step: 220, loss: 0.04041646048426628
step: 230, loss: 0.04373852163553238
step: 240, loss: 0.0003853653615806252
step: 250, loss: 0.00020224342006258667
step: 260, loss: 0.0011891323374584317
step: 270, loss: 0.00029534738860093057
step: 280, loss: 0.001407568110153079
step: 290, loss: 0.00012113218690501526
step: 300, loss: 0.0005571968504227698
step: 310, loss: 0.00012381187116261572
step: 320, loss: 0.0007892398862168193
step: 330, loss: 0.0003750785253942013
step: 340, loss: 0.0012197927571833134
step: 350, loss: 0.006476098205894232
step: 360, loss: 0.00012498689466156065
step: 370, loss: 0.0010144597617909312
step: 380, loss: 0.025364622473716736
step: 390, loss: 0.009085563011467457
step: 400, loss: 0.0004309635842218995
step: 410, loss: 0.008837143890559673
step: 420, loss: 0.005555488634854555
step: 430, loss: 0.0003876538830809295
step: 440, loss: 0.007549551781266928
step: 450, loss: 0.0003563159261830151
step: 460, loss: 0.00015309915761463344
step: 470, loss: 0.00015895810793153942
step: 480, loss: 0.0027588484808802605
step: 490, loss: 0.0005773674347437918
step: 500, loss: 0.07050382345914841
step: 510, loss: 0.0001698032720014453
step: 520, loss: 0.0001550184824736789
step: 530, loss: 0.0002936673117801547
/usr/local/lib/python3.7/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=0.913556920170052, f1=0.89058039961941,
best_f1=0.889947594092425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02111021988093853
step: 10, loss: 0.0011768217664211988
step: 20, loss: 0.0001976555649889633
step: 30, loss: 6.44249957986176e-05
step: 40, loss: 0.00011734857980627567
```

step: 50, loss: 0.00014975109661463648
step: 60, loss: 0.0008422690443694592
step: 70, loss: 8.032030746107921e-05
step: 80, loss: 0.0010406498331576586
step: 90, loss: 0.0004944023094139993
step: 100, loss: 0.0010081232758238912
step: 110, loss: 0.00011938439274672419
step: 120, loss: 0.0003095013089478016
step: 130, loss: 6.270306039368734e-05
step: 140, loss: 7.885642116889358e-05
step: 150, loss: 8.540981798432767e-05
step: 160, loss: 0.00011241518222959712
step: 170, loss: 4.992314279661514e-05
step: 180, loss: 0.0006235938635654747
step: 190, loss: 0.0005781258805654943
step: 200, loss: 0.003053235588595271
step: 210, loss: 0.0028142656665295362
step: 220, loss: 4.143767000641674e-05
step: 230, loss: 2.5730125344125554e-05
step: 240, loss: 0.00012899631110485643
step: 250, loss: 0.00036957176052965224
step: 260, loss: 8.68267088662833e-05
step: 270, loss: 0.00015749363228678703
step: 280, loss: 0.0003287859435658902
step: 290, loss: 0.0001972472236957401
step: 300, loss: 0.00014471652684733272
step: 310, loss: 0.00010136430501006544
step: 320, loss: 0.00010350941010983661
step: 330, loss: 0.0009598037577234209
step: 340, loss: 7.830112735973671e-05
step: 350, loss: 7.39971874281764e-05
step: 360, loss: 6.011468212818727e-05
step: 370, loss: 0.0003026325721293688
step: 380, loss: 4.720401193480939e-05
step: 390, loss: 3.690996891236864e-05
step: 400, loss: 2.7275866159470752e-05
step: 410, loss: 0.00011902152618858963
step: 420, loss: 0.00012295150372665375
step: 430, loss: 9.190943819703534e-05
step: 440, loss: 0.00012096231512259692
step: 450, loss: 0.00021136719442438334
step: 460, loss: 0.001286514918319881
step: 470, loss: 3.9702463254798204e-05
step: 480, loss: 0.00013937879703007638
step: 490, loss: 5.70999582123477e-05
step: 500, loss: 0.0011171390069648623
step: 510, loss: 5.084829172119498e-05
step: 520, loss: 0.00014991391799412668

```
step: 530, loss: 6.31698640063405e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9135460009246417, f1=0.8953703703703704,
best_f1=0.889947594092425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003404116432648152
step: 10, loss: 4.9549664254300296e-05
step: 20, loss: 0.00023950946342665702
step: 30, loss: 0.00014367714175023139
step: 40, loss: 0.00010599356755847111
step: 50, loss: 0.001389363664202392
step: 60, loss: 0.00020533811766654253
step: 70, loss: 8.835098560666665e-05
step: 80, loss: 0.00021774330525659025
step: 90, loss: 8.097830141196027e-05
step: 100, loss: 0.00011417666974011809
step: 110, loss: 7.953298336360604e-05
step: 120, loss: 2.9235079637146555e-05
step: 130, loss: 3.0170356694725342e-05
step: 140, loss: 6.5217463998124e-05
step: 150, loss: 0.0001035171007970348
step: 160, loss: 4.005975279142149e-05
step: 170, loss: 0.00038689534994773567
step: 180, loss: 5.4372438171412796e-05
step: 190, loss: 0.00010930919961538166
step: 200, loss: 6.572675192728639e-05
step: 210, loss: 9.856825636234134e-05
step: 220, loss: 0.00022331053332891315
step: 230, loss: 4.424508006195538e-05
step: 240, loss: 5.003048136131838e-05
step: 250, loss: 6.867313641123474e-05
step: 260, loss: 3.0211218472686596e-05
step: 270, loss: 8.307973621413112e-05
step: 280, loss: 0.0008166661718860269
step: 290, loss: 0.0006354317883960903
step: 300, loss: 0.001995209837332368
step: 310, loss: 5.706663796445355e-05
step: 320, loss: 5.478324965224601e-05
step: 330, loss: 4.534568506642245e-05
step: 340, loss: 0.0002329457929590717
step: 350, loss: 0.0001000734482659027
step: 360, loss: 0.002868547337129712
step: 370, loss: 0.001583907869644463
step: 380, loss: 8.201748278224841e-05
```

```

step: 390, loss: 0.0028158584609627724
step: 400, loss: 3.3592747058719397e-05
step: 410, loss: 9.977760055335239e-05
step: 420, loss: 0.000412529829191044
step: 430, loss: 0.00011311750131426379
step: 440, loss: 0.0038392001297324896
step: 450, loss: 0.002595327328890562
step: 460, loss: 2.6556963348411955e-05
step: 470, loss: 4.7064913815120235e-05
step: 480, loss: 0.00045280990889295936
step: 490, loss: 0.0031165245454758406
step: 500, loss: 4.558470391202718e-05
step: 510, loss: 0.0003561972116585821
step: 520, loss: 3.70800044038333e-05
step: 530, loss: 0.0004871632845606655
/usr/local/lib/python3.7/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=0.9042553191489362, f1=0.886100386100386,
best_f1=0.889947594092425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015745146083645523
step: 10, loss: 9.304402919951826e-05
step: 20, loss: 3.381955320946872e-05
step: 30, loss: 0.0007307616760954261
step: 40, loss: 3.5481032682582736e-05
step: 50, loss: 0.0014686435461044312
step: 60, loss: 5.746229726355523e-05
step: 70, loss: 0.02731066197156906
step: 80, loss: 2.8038935852237046e-05
step: 90, loss: 8.963628351921216e-05
step: 100, loss: 6.49827197776176e-05
step: 110, loss: 2.9030334189883433e-05
step: 120, loss: 4.551086749415845e-05
step: 130, loss: 2.3830169084249064e-05
step: 140, loss: 4.227560930303298e-05
step: 150, loss: 0.0014924018178135157
step: 160, loss: 0.0005593406967818737
step: 170, loss: 0.00025444693164899945
step: 180, loss: 0.0011282551568001509
step: 190, loss: 0.0017732041887938976
step: 200, loss: 0.0006270114681683481
step: 210, loss: 0.00012371221964713186
step: 220, loss: 0.000161341144121252
step: 230, loss: 0.024482646957039833
step: 240, loss: 7.695711246924475e-05

```



```

step: 250, loss: 0.00022634475317317992
step: 260, loss: 8.777147741056979e-05
step: 270, loss: 0.0005056242225691676
step: 280, loss: 4.3806765461340547e-05
step: 290, loss: 0.0016594841144979
step: 300, loss: 3.29529975715559e-05
step: 310, loss: 0.00022640109818894416
step: 320, loss: 5.7157845731126145e-05
step: 330, loss: 0.00026171133504249156
step: 340, loss: 6.714831397403032e-05
step: 350, loss: 9.201405919156969e-05
step: 360, loss: 5.1146063924534246e-05
step: 370, loss: 0.06446170061826706
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 380, loss: 0.0003439753782004118
step: 390, loss: 0.002040534047409892
step: 400, loss: 0.0006849839701317251
step: 410, loss: 0.0014939485117793083
step: 420, loss: 0.0001298888528253883
step: 430, loss: 0.00017974649381358176
step: 440, loss: 0.0010405133944004774
step: 450, loss: 0.00037506394437514246
step: 460, loss: 0.004927139729261398
step: 470, loss: 3.5272296372568235e-05
step: 480, loss: 3.7232272006804124e-05
step: 490, loss: 7.956875197123736e-05
step: 500, loss: 2.4141785615938716e-05
step: 510, loss: 0.0001354128762613982
step: 520, loss: 0.004116499796509743
step: 530, loss: 3.056776768062264e-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=0.9166269652215341, f1=0.8964525407478429,
best_f1=0.889947594092425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001526606356492266
step: 10, loss: 2.2864855054649524e-05
step: 20, loss: 4.016093225800432e-05
step: 30, loss: 4.5231241529108956e-05
step: 40, loss: 0.00468703405931592
step: 50, loss: 0.00012616468302439898
step: 60, loss: 0.00023621792206540704
step: 70, loss: 0.00011371474101906642
step: 80, loss: 4.729598731501028e-05
step: 90, loss: 0.00031319676782004535

```

```
step: 100, loss: 6.455371476477012e-05
step: 110, loss: 7.781365275150165e-05
step: 120, loss: 3.925488636014052e-05
step: 130, loss: 0.0003843719605356455
step: 140, loss: 3.105607538600452e-05
step: 150, loss: 8.650263043818995e-05
step: 160, loss: 4.015272133983672e-05
step: 170, loss: 0.0001378385495627299
step: 180, loss: 7.402622577501461e-05
step: 190, loss: 0.0002154657559003681
step: 200, loss: 6.418829434551299e-05
step: 210, loss: 2.5361105144838803e-05
step: 220, loss: 0.00026888586580753326
step: 230, loss: 0.0001845984807005152
step: 240, loss: 3.4728778700809926e-05
step: 250, loss: 4.995870403945446e-05
step: 260, loss: 6.69581422698684e-05
step: 270, loss: 3.3660235203569755e-05
step: 280, loss: 6.152755668153986e-05
step: 290, loss: 0.0004258162807673216
step: 300, loss: 0.001284162630327046
step: 310, loss: 0.00039667863165959716
step: 320, loss: 0.0001371239632135257
step: 330, loss: 0.00018212666327599436
step: 340, loss: 5.159999636816792e-05
step: 350, loss: 0.00012943956244271249
step: 360, loss: 0.00037842255551368
step: 370, loss: 4.1345181671204045e-05
step: 380, loss: 8.012894977582619e-05
step: 390, loss: 0.0013282395666465163
step: 400, loss: 0.00026169285411015153
step: 410, loss: 7.041277422104031e-05
step: 420, loss: 3.37548553943634e-05
step: 430, loss: 7.602661935379729e-05
step: 440, loss: 0.0007081864168867469
step: 450, loss: 0.0002722225326579064
step: 460, loss: 3.916819332516752e-05
step: 470, loss: 3.835623647319153e-05
step: 480, loss: 0.00011851538874907419
step: 490, loss: 3.401615686016157e-05
step: 500, loss: 1.938965215231292e-05
step: 510, loss: 0.00024447261239401996
step: 520, loss: 7.824298518244177e-05
step: 530, loss: 3.270643355790526e-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=0.9189957366177167, f1=0.8956646021915198,

best_f1=0.8956646021915198

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

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

step: 0, loss: 4.069894930580631e-05

step: 10, loss: 0.000161455842317082

step: 20, loss: 8.132925722748041e-05

step: 30, loss: 0.00025478930911049247

step: 40, loss: 8.377403719350696e-05

step: 50, loss: 0.00017424052930437028

step: 60, loss: 1.8622364223119803e-05

step: 70, loss: 0.0002314460143679753

step: 80, loss: 0.0004985517007298768

step: 90, loss: 6.481949822045863e-05

step: 100, loss: 7.066282705636695e-05

step: 110, loss: 5.032658009440638e-05

step: 120, loss: 2.9495780836441554e-05

step: 130, loss: 0.0013322783634066582

step: 140, loss: 0.00014256307622417808

step: 150, loss: 0.00015319319209083915

step: 160, loss: 8.690202957950532e-05

step: 170, loss: 0.0006415120442397892

step: 180, loss: 2.2600666852667928e-05

step: 190, loss: 0.0005030771135352552

step: 200, loss: 2.828879951266572e-05

step: 210, loss: 9.353106725029647e-05

step: 220, loss: 5.293632057146169e-05

step: 230, loss: 5.883479389012791e-05

step: 240, loss: 0.0040060714818537235

step: 250, loss: 3.836063842754811e-05

step: 260, loss: 2.766690704447683e-05

step: 270, loss: 3.165623274981044e-05

step: 280, loss: 0.0001830161054385826

step: 290, loss: 2.494380714779254e-05

step: 300, loss: 4.4052543671568856e-05

step: 310, loss: 0.00039586317143402994

step: 320, loss: 0.005367023404687643

step: 330, loss: 6.27185363555327e-05

step: 340, loss: 2.6381669158581644e-05

step: 350, loss: 1.640597838559188e-05

step: 360, loss: 0.0015970732783898711

step: 370, loss: 0.0020748034585267305

step: 380, loss: 3.647966877906583e-05

step: 390, loss: 5.453384073916823e-05

```

step: 400, loss: 3.4047938243020326e-05
step: 410, loss: 0.00012536977010313421
step: 420, loss: 0.00014161197759676725
step: 430, loss: 0.0001265190076082945
step: 440, loss: 5.0983624532818794e-05
step: 450, loss: 2.6351588530815206e-05
step: 460, loss: 3.998633474111557e-05
step: 470, loss: 2.9401928259176202e-05
step: 480, loss: 4.06507751904428e-05
step: 490, loss: 2.2150126824271865e-05
step: 500, loss: 3.170039417454973e-05
step: 510, loss: 0.00012333244376350194
step: 520, loss: 3.890774678438902e-05
step: 530, loss: 3.142894274787977e-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=0.9203956665096561, f1=0.8980169971671387,
best_f1=0.8980169971671387
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:24, 232.26it/s]
load_f1 = 0.9186155285313378
real_f1 = 0.9163533834586467
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 242.50it/s]

```

4 DITTO

4.1 DITTO STRUCTURED

4.1.1 Beer - Running the matcher

```

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

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

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

```

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

step: 0, loss: 0.5551786422729492

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

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

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37333333333333335, f1=0.33766233766233766,
best_f1=0.33766233766233766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5203631520271301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.42424242424242425, f1=0.3835616438356164,
best_f1=0.3835616438356164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5861638784408569
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5306122448979592, f1=0.3934426229508196,
best_f1=0.3934426229508196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3425706923007965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5777777777777778, 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: 0.36081892251968384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6341463414634146, f1=0.4878048780487805,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29483652114868164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7142857142857143, f1=0.5806451612903226,
best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2356729656457901
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7272727272727273, f1=0.55, best_f1=0.55
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05478467792272568
/usr/local/lib/python3.7/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=0.7142857142857143, f1=0.5625000000000001, best_f1=0.55
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00521625904366374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7647058823529412, f1=0.6111111111111112,
best_f1=0.6111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.062225230038166046
/usr/local/lib/python3.7/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=0.7027027027027025, f1=0.5641025641025641,
best_f1=0.6111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009794116951525211
/usr/local/lib/python3.7/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=0.7200000000000001, f1=0.5517241379310344,
best_f1=0.6111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05906094238162041
/usr/local/lib/python3.7/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=0.742857142857143, f1=0.5641025641025641,
best_f1=0.6111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020768988877534866
/usr/local/lib/python3.7/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=0.742857142857143, f1=0.5641025641025641,
best_f1=0.6111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013823959976434708

```

```

/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.5641025641025641,
best_f1=0.6111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030742648988962173
/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.5641025641025641,
best_f1=0.6111111111111112
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 138642.09it/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.")
load_f1 = 0.6923076923076924
real_f1 = 0.6923076923076924
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:17, 247.00it/s]
```

4.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6125273108482361
step: 10, loss: 0.634732723236084
step: 20, loss: 0.48264914751052856
step: 30, loss: 0.37067699432373047
step: 40, loss: 0.15597350895404816
step: 50, loss: 0.33545631170272827
step: 60, loss: 0.04662356525659561
step: 70, loss: 0.024408916011452675
step: 80, loss: 0.13388077914714813
step: 90, loss: 0.21603842079639435
step: 100, loss: 0.01286372635513544
step: 110, loss: 0.1714010387659073
step: 120, loss: 0.03181290999054909
step: 130, loss: 0.0228571817278862
step: 140, loss: 0.0061898622661828995
step: 150, loss: 0.08743108063936234
step: 160, loss: 0.011198822408914566
step: 170, loss: 0.024940330535173416

```

```

step: 180, loss: 0.1438705027103424
step: 190, loss: 0.14274956285953522
step: 200, loss: 0.02236882410943508
step: 210, loss: 0.004620740655809641
step: 220, loss: 0.06944581121206284
step: 230, loss: 0.048176731914281845
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.963882618510158, f1=0.9625425652667423,
best_f1=0.9625425652667423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03449138626456261
step: 10, loss: 0.0035242901649326086
step: 20, loss: 0.07605284452438354
step: 30, loss: 0.19910404086112976
step: 40, loss: 0.017395203933119774
step: 50, loss: 0.006793178617954254
step: 60, loss: 0.004923023283481598
step: 70, loss: 0.15433919429779053
step: 80, loss: 0.022104522213339806
step: 90, loss: 0.016230657696723938
step: 100, loss: 0.0994827151298523
step: 110, loss: 0.06646080315113068
step: 120, loss: 0.031216736882925034
step: 130, loss: 0.020358191803097725
step: 140, loss: 0.03473959118127823
step: 150, loss: 0.011558104306459427
step: 160, loss: 0.0028204198461025953
step: 170, loss: 0.022912971675395966
step: 180, loss: 0.014498892240226269
step: 190, loss: 0.017853576689958572
step: 200, loss: 0.002065579406917095
step: 210, loss: 0.0008340222993865609
step: 220, loss: 0.09280554950237274
step: 230, loss: 0.09112107008695602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9730337078651685,
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: 0.008639758452773094
step: 10, loss: 0.003373668296262622
step: 20, loss: 0.0072344038635492325
step: 30, loss: 0.003261468606069684
step: 40, loss: 0.0993417352437973
step: 50, loss: 0.04753194376826286
step: 60, loss: 0.004276916850358248
step: 70, loss: 0.10581675916910172
step: 80, loss: 0.0008523485739715397
step: 90, loss: 0.1360769271850586
step: 100, loss: 0.0012634432641789317
step: 110, loss: 0.0018622373463585973
step: 120, loss: 0.016055045649409294
step: 130, loss: 0.0011884871637448668
step: 140, loss: 0.058881524950265884
step: 150, loss: 0.025226986035704613
step: 160, loss: 0.0019915227312594652
step: 170, loss: 0.00471445033326745
step: 180, loss: 0.023233307525515556
step: 190, loss: 0.019532231613993645
step: 200, loss: 0.040428485721349716
step: 210, loss: 0.0013366129714995623
step: 220, loss: 0.0012954672565683722
step: 230, loss: 0.0010143450926989317
/usr/local/lib/python3.7/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=0.9754464285714286, f1=0.9732142857142857,
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: 0.0013093689922243357
step: 10, loss: 0.0003755372017621994
step: 20, loss: 0.004208176862448454
step: 30, loss: 0.009115096181631088
step: 40, loss: 0.01641671359539032
step: 50, loss: 0.005156326573342085
step: 60, loss: 0.0395493283867836
step: 70, loss: 0.00042582247988320887
step: 80, loss: 0.009618241339921951
step: 90, loss: 0.0008346524555236101

```

```

step: 100, loss: 0.0010335530387237668
step: 110, loss: 0.001435709884390235
step: 120, loss: 0.03737922012805939
step: 130, loss: 0.0024665528908371925
step: 140, loss: 0.027104660868644714
step: 150, loss: 0.10361763834953308
step: 160, loss: 0.004091465380042791
step: 170, loss: 0.17546454071998596
step: 180, loss: 0.0007796854479238391
step: 190, loss: 0.01473039947450161
step: 200, loss: 0.0009342360426671803
step: 210, loss: 0.09205443412065506
step: 220, loss: 0.0013719445560127497
step: 230, loss: 0.0020995272789150476
/usr/local/lib/python3.7/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=0.9757709251101322, f1=0.9649890590809628,
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: 0.0011607721680775285
step: 10, loss: 0.00782892107963562
step: 20, loss: 0.037098679691553116
step: 30, loss: 0.00045101146679371595
step: 40, loss: 0.0002659203018993139
step: 50, loss: 0.00047101982636377215
step: 60, loss: 0.0007552462047897279
step: 70, loss: 0.00019178990623913705
step: 80, loss: 0.001598771894350648
step: 90, loss: 0.0003008981002494693
step: 100, loss: 0.0008166742045432329
step: 110, loss: 0.00019008545496035367
step: 120, loss: 0.00037482636980712414
step: 130, loss: 0.009593593887984753
step: 140, loss: 0.008788051083683968
step: 150, loss: 0.015309691429138184
step: 160, loss: 0.0037202350795269012
step: 170, loss: 0.003069077618420124
step: 180, loss: 0.11289866268634796
step: 190, loss: 0.0005789376446045935
step: 200, loss: 0.04977957531809807
step: 210, loss: 0.0033428766764700413
step: 220, loss: 0.002379925223067403
step: 230, loss: 0.0002483751450199634
/usr/local/lib/python3.7/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=0.9730337078651685, f1=0.9777777777777777,
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: 0.001607892569154501
step: 10, loss: 0.010885201394557953
step: 20, loss: 0.004166122525930405
step: 30, loss: 0.0010889994446188211
step: 40, loss: 0.001360714202746749
step: 50, loss: 0.03767545893788338
step: 60, loss: 0.00019519656780175865
step: 70, loss: 0.0008439291850663722
step: 80, loss: 0.0009751093457452953
step: 90, loss: 0.006544723175466061
step: 100, loss: 0.030971093103289604
step: 110, loss: 0.04281361401081085
step: 120, loss: 0.0004070907016284764
step: 130, loss: 0.008809477090835571
step: 140, loss: 0.0075585199519991875
step: 150, loss: 0.0013264210429042578
step: 160, loss: 0.014085362665355206
step: 170, loss: 0.0040975287556648254
step: 180, loss: 0.05494081974029541
step: 190, loss: 0.0018045371398329735
step: 200, loss: 0.0011145866010338068
step: 210, loss: 0.0017107243184000254
step: 220, loss: 0.0004072347073815763
step: 230, loss: 0.029694009572267532
/usr/local/lib/python3.7/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=0.9763779527559054, f1=0.9751693002257337,
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: 0.019787998870015144
step: 10, loss: 9.946034697350115e-05
step: 20, loss: 0.00025518020265735686
step: 30, loss: 0.0002594663528725505
step: 40, loss: 0.00032137450762093067
step: 50, loss: 0.0003699635562952608
step: 60, loss: 0.000748339225538075
step: 70, loss: 0.030943134799599648
step: 80, loss: 0.0005708155804313719
step: 90, loss: 0.0012621150817722082

```



```

step: 100, loss: 0.000265562062850222
step: 110, loss: 0.0010227069724351168
step: 120, loss: 0.0006022921297699213
step: 130, loss: 0.0008827627752907574
step: 140, loss: 0.005129074677824974
step: 150, loss: 0.001647222088649869
step: 160, loss: 0.09777458757162094
step: 170, loss: 0.0008839906076900661
step: 180, loss: 0.0003091710968874395
step: 190, loss: 0.0018328444566577673
step: 200, loss: 0.06617086380720139
step: 210, loss: 0.0004392415867187083
step: 220, loss: 0.0005553118535317481
step: 230, loss: 0.001644135918468237
/usr/local/lib/python3.7/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=0.9729119638826186, f1=0.9683972911963882,
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: 0.000585924310144037
step: 10, loss: 0.021533535793423653
step: 20, loss: 0.00014904244744684547
step: 30, loss: 0.00015236393664963543
step: 40, loss: 0.0026699211448431015
step: 50, loss: 0.0008456922369077802
step: 60, loss: 0.0101921446621418
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.0004056961042806506
step: 80, loss: 0.0002688123204279691
step: 90, loss: 8.046348375501111e-05
step: 100, loss: 0.0004361363244242966
step: 110, loss: 0.00015628825349267572
step: 120, loss: 0.04338430240750313
step: 130, loss: 0.000979208038188517
step: 140, loss: 0.000273028650553897
step: 150, loss: 0.011321723461151123
step: 160, loss: 0.0001232421345775947
step: 170, loss: 0.0002896888763643801
step: 180, loss: 0.0006927576032467186
step: 190, loss: 0.00023239188885781914
step: 200, loss: 0.00014466750144492835
step: 210, loss: 0.00013214141654316336
step: 220, loss: 0.00012104208872187883
step: 230, loss: 0.00013071011926513165
/usr/local/lib/python3.7/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=0.9743589743589743, f1=0.9722530521642618,
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: 0.00018082499445881695
step: 10, loss: 0.00016993182362057269
step: 20, loss: 0.0007032356807030737
step: 30, loss: 3.9307251427089795e-05
step: 40, loss: 0.023064885288476944
step: 50, loss: 5.013382178731263e-05
step: 60, loss: 9.287639113608748e-05
step: 70, loss: 0.0005080051487311721
step: 80, loss: 0.0006531900144182146
step: 90, loss: 7.249110785778612e-05
step: 100, loss: 0.00021977751748636365
step: 110, loss: 8.527059253538027e-05
step: 120, loss: 3.847735933959484e-05
step: 130, loss: 0.001056784181855619
step: 140, loss: 0.0003156095917802304
step: 150, loss: 0.06459996849298477
step: 160, loss: 0.00043831937364302576
step: 170, loss: 0.0002595716214273125
step: 180, loss: 0.007869930937886238
step: 190, loss: 0.00042988997302018106
step: 200, loss: 0.0049122171476483345
step: 210, loss: 0.0032553470227867365
step: 220, loss: 0.00019230431644245982
step: 230, loss: 0.0004990277229808271
/usr/local/lib/python3.7/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=0.9680365296803655, f1=0.9681093394077448,
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: 0.0028655147179961205
step: 10, loss: 0.0001326439087279141
step: 20, loss: 0.0012397492537274957
step: 30, loss: 0.00030531291849911213
step: 40, loss: 0.00012941742897965014
step: 50, loss: 0.00050439836923033
step: 60, loss: 0.003960700239986181
step: 70, loss: 0.0012088401708751917
step: 80, loss: 8.564864401705563e-05

```

```

step: 90, loss: 0.0036175663117319345
step: 100, loss: 0.00021624584042001516
step: 110, loss: 0.002810389269143343
step: 120, loss: 0.0003510891692712903
step: 130, loss: 0.00036297127371653914
step: 140, loss: 0.05609610676765442
step: 150, loss: 0.00908343493938446
step: 160, loss: 0.0099077969789505
step: 170, loss: 5.8438330597709864e-05
step: 180, loss: 0.0007637699018232524
step: 190, loss: 0.08078867942094803
step: 200, loss: 0.00043403799645602703
step: 210, loss: 0.00022183007968124002
step: 220, loss: 0.0004024274239782244
step: 230, loss: 0.038544390350580215
/usr/local/lib/python3.7/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=0.9730941704035874, f1=0.9743016759776536,
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: 9.082709584617987e-05
step: 10, loss: 0.00034815858816727996
step: 20, loss: 0.009160144254565239
step: 30, loss: 0.007618098519742489
step: 40, loss: 0.0010197034571319818
step: 50, loss: 0.000751265324652195
step: 60, loss: 0.0004319300933275372
step: 70, loss: 0.024562230333685875
step: 80, loss: 0.0007184858550317585
step: 90, loss: 7.791447569616139e-05
step: 100, loss: 0.00010775140253826976
step: 110, loss: 0.0002952012000605464
step: 120, loss: 0.000107263978861738
step: 130, loss: 0.0003135384467896074
step: 140, loss: 0.0021705098915845156
step: 150, loss: 0.006647143047302961
step: 160, loss: 8.060748950811103e-05
step: 170, loss: 0.0008701553451828659
step: 180, loss: 0.008721905760467052
step: 190, loss: 4.3743508285842836e-05
step: 200, loss: 7.770375668769702e-05
step: 210, loss: 7.78487155912444e-05
step: 220, loss: 0.0001951821323018521
step: 230, loss: 0.0007567350403405726
/usr/local/lib/python3.7/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=0.9730337078651685, f1=0.9752252252252253,
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: 4.0578153857495636e-05
step: 10, loss: 4.968493522028439e-05
step: 20, loss: 4.513595558819361e-05
step: 30, loss: 9.369882172904909e-05
step: 40, loss: 0.0020162700675427914
step: 50, loss: 0.005558932665735483
step: 60, loss: 0.0017953802598640323
step: 70, loss: 8.840008376864716e-05
step: 80, loss: 0.00010667923925211653
step: 90, loss: 3.569846012396738e-05
step: 100, loss: 0.0002167539350921288
step: 110, loss: 6.661543739028275e-05
step: 120, loss: 0.0001890489220386371
step: 130, loss: 9.409159974893555e-05
step: 140, loss: 7.547594577772543e-05
step: 150, loss: 0.0028453092090785503
step: 160, loss: 0.00021284635295160115
step: 170, loss: 7.920097414171323e-05
step: 180, loss: 0.00011355476453900337
step: 190, loss: 6.694899639114738e-05
step: 200, loss: 2.4880751880118623e-05
step: 210, loss: 4.119645745959133e-05
step: 220, loss: 0.002501735696569085
step: 230, loss: 0.0030976622365415096
/usr/local/lib/python3.7/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=0.9743589743589743, f1=0.9744160177975528,
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: 0.0019896957091987133
step: 10, loss: 0.00018761609680950642
step: 20, loss: 0.0022615455090999603
step: 30, loss: 0.0011208084179088473
step: 40, loss: 0.12012497335672379
step: 50, loss: 9.66516017797403e-05
step: 60, loss: 3.823845327133313e-05
step: 70, loss: 3.387539982213639e-05
step: 80, loss: 3.867389386869036e-05

```

```

step: 90, loss: 0.0004848790413234383
step: 100, loss: 4.9339225370204076e-05
step: 110, loss: 0.005670326761901379
step: 120, loss: 8.095484372461215e-05
step: 130, loss: 8.522148709744215e-05
step: 140, loss: 3.758263846975751e-05
step: 150, loss: 2.836006933648605e-05
step: 160, loss: 0.00038166638114489615
step: 170, loss: 6.332749035209417e-05
step: 180, loss: 3.661058144643903e-05
step: 190, loss: 4.059881757711992e-05
step: 200, loss: 0.0030685788951814175
step: 210, loss: 2.451550608384423e-05
step: 220, loss: 7.298430864466354e-05
step: 230, loss: 3.574219226720743e-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=0.9764837625979844, f1=0.9764309764309763,
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: 4.207823803881183e-05
step: 10, loss: 0.01126464270055294
step: 20, loss: 0.0001288648636545986
step: 30, loss: 0.0006197839393280447
step: 40, loss: 0.0005366250406950712
step: 50, loss: 5.945978773524985e-05
step: 60, loss: 4.061816798639484e-05
step: 70, loss: 0.00016805788618512452
step: 80, loss: 0.0005472922930493951
step: 90, loss: 3.449474388617091e-05
step: 100, loss: 4.3041804019594565e-05
step: 110, loss: 3.908125654561445e-05
step: 120, loss: 2.2910164261702448e-05
step: 130, loss: 3.299772652098909e-05
step: 140, loss: 4.0667018765816465e-05
step: 150, loss: 2.7462168873171322e-05
step: 160, loss: 0.005897959228605032
step: 170, loss: 2.3290140234166756e-05
step: 180, loss: 4.888570765615441e-05
step: 190, loss: 3.502632534946315e-05
step: 200, loss: 4.819318201043643e-05
step: 210, loss: 0.00028195496997796
step: 220, loss: 0.0007639891118742526
step: 230, loss: 0.00029167826869525015
/usr/local/lib/python3.7/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=0.9751693002257337, f1=0.9763779527559054,
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: 3.6825815186602995e-05
step: 10, loss: 2.7737132768379524e-05
step: 20, loss: 6.162388308439404e-05
step: 30, loss: 4.683159932028502e-05
step: 40, loss: 0.0002474795328453183
step: 50, loss: 0.02606639824807644
step: 60, loss: 2.9306022042874247e-05
step: 70, loss: 8.063529821811244e-05
step: 80, loss: 0.0097003523260355
step: 90, loss: 8.505635923938826e-05
step: 100, loss: 5.371090810513124e-05
step: 110, loss: 0.0048848786391317844
step: 120, loss: 0.00021831806225236505
step: 130, loss: 2.91050091618672e-05
step: 140, loss: 2.7816096917376854e-05
step: 150, loss: 6.428708002204075e-05
step: 160, loss: 0.0001923644740600139
step: 170, loss: 0.00010464182560099289
step: 180, loss: 4.6916858991608024e-05
step: 190, loss: 0.00012126524961786345
step: 200, loss: 8.638860163046047e-05
step: 210, loss: 0.00014564265438821167
step: 220, loss: 0.0006828467594459653
step: 230, loss: 0.0001540316443424672
/usr/local/lib/python3.7/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=0.976324689966178, f1=0.9763779527559054,
best_f1=0.9730337078651685
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 193.12it/s]
load_f1 = 0.9797297297297298
real_f1 = 0.9808773903262092
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 234.74it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--da attr_del

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

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

```

```

Downloading: 100% 433/433 [00:00<00:00, 395kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 347kB/s]
Downloading: 100% 440M/440M [00:24<00:00, 17.7MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

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

```

```

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

```

```

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: 0.6516682505607605
step: 10, loss: 0.5442603230476379
step: 20, loss: 0.5805392265319824
step: 30, loss: 0.354943186044693
step: 40, loss: 0.43120574951171875
step: 50, loss: 0.3120923638343811
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.09215760976076126
step: 70, loss: 0.27884456515312195
step: 80, loss: 0.0733758956193924
step: 90, loss: 0.47419053316116333
step: 100, loss: 0.10777720808982849
step: 110, loss: 0.09580127149820328
step: 120, loss: 0.1123201847076416
step: 130, loss: 0.12251872569322586
step: 140, loss: 0.15984408557415009
step: 150, loss: 0.13865146040916443
step: 160, loss: 0.04552578181028366
step: 170, loss: 0.25079792737960815
step: 180, loss: 0.05609835311770439
step: 190, loss: 0.05890258774161339
step: 200, loss: 0.13212767243385315
step: 210, loss: 0.0864730030298233
step: 220, loss: 0.325448602437973
step: 230, loss: 0.15821203589439392
step: 240, loss: 0.09702880680561066
step: 250, loss: 0.11176594346761703
step: 260, loss: 0.26485079526901245
step: 270, loss: 0.027711452916264534
step: 280, loss: 0.11221151053905487
step: 290, loss: 0.1523686647415161
step: 300, loss: 0.053268443793058395
step: 310, loss: 0.17913196980953217
step: 320, loss: 0.0817926675081253
step: 330, loss: 0.045366544276475906
step: 340, loss: 0.12988236546516418
step: 350, loss: 0.034080374985933304
step: 360, loss: 0.05461651459336281
step: 370, loss: 0.21469518542289734
step: 380, loss: 0.020308375358581543
```

```

step: 390, loss: 0.37779682874679565
step: 400, loss: 0.22145172953605652
step: 410, loss: 0.0811118483543396
step: 420, loss: 0.15984764695167542
step: 430, loss: 0.14339867234230042
step: 440, loss: 0.04561661556363106
step: 450, loss: 0.029530050233006477
step: 460, loss: 0.019099729135632515
step: 470, loss: 0.09894194453954697
step: 480, loss: 0.12412215024232864
step: 490, loss: 0.24672026932239532
step: 500, loss: 0.13427023589611053
step: 510, loss: 0.0851416140794754
step: 520, loss: 0.04163914918899536
step: 530, loss: 0.036168716847896576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9001865671641792, f1=0.8962308050255934,
best_f1=0.8962308050255934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18727123737335205
step: 10, loss: 0.12514831125736237
step: 20, loss: 0.03884948045015335
step: 30, loss: 0.028027797117829323
step: 40, loss: 0.11921899020671844
step: 50, loss: 0.1521671712398529
step: 60, loss: 0.038016147911548615
step: 70, loss: 0.08387494087219238
step: 80, loss: 0.22234179079532623
step: 90, loss: 0.034747861325740814
step: 100, loss: 0.12771742045879364
step: 110, loss: 0.07001998275518417
step: 120, loss: 0.13377292454242706
step: 130, loss: 0.15545545518398285
step: 140, loss: 0.0590069517493248
step: 150, loss: 0.13714823126792908
step: 160, loss: 0.021254075691103935
step: 170, loss: 0.0721975788474083
step: 180, loss: 0.03301473706960678
step: 190, loss: 0.05148835852742195
step: 200, loss: 0.023888645693659782

```

```

step: 210, loss: 0.0598914809525013
step: 220, loss: 0.08789756149053574
step: 230, loss: 0.009201704524457455
step: 240, loss: 0.042390793561935425
step: 250, loss: 0.07698938995599747
step: 260, loss: 0.01227183174341917
step: 270, loss: 0.2684674859046936
step: 280, loss: 0.07074382901191711
step: 290, loss: 0.08378026634454727
step: 300, loss: 0.03767503425478935
step: 310, loss: 0.11072684079408646
step: 320, loss: 0.19647139310836792
step: 330, loss: 0.06642482429742813
step: 340, loss: 0.043532196432352066
step: 350, loss: 0.004251889418810606
step: 360, loss: 0.06392479687929153
step: 370, loss: 0.1327272206544876
step: 380, loss: 0.19990655779838562
step: 390, loss: 0.1615966409444809
step: 400, loss: 0.04357128217816353
step: 410, loss: 0.010058870539069176
step: 420, loss: 0.062362439930438995
step: 430, loss: 0.035506147891283035
step: 440, loss: 0.045352593064308167
step: 450, loss: 0.0999252200126648
step: 460, loss: 0.043159130960702896
step: 470, loss: 0.013273250311613083
step: 480, loss: 0.1657121628522873
step: 490, loss: 0.02465713582932949
step: 500, loss: 0.30680060386657715
step: 510, loss: 0.04114522039890289
step: 520, loss: 0.12404020875692368
step: 530, loss: 0.023780034855008125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9096045197740114, f1=0.9001865671641792,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08650670945644379
step: 10, loss: 0.044175148010253906
step: 20, loss: 0.1772846281528473

```

step: 30, loss: 0.05118848755955696
step: 40, loss: 0.03791843354701996
step: 50, loss: 0.13140372931957245
step: 60, loss: 0.048562247306108475
step: 70, loss: 0.011762753129005432
step: 80, loss: 0.00736711360514164
step: 90, loss: 0.02990984544157982
step: 100, loss: 0.07369312644004822
step: 110, loss: 0.0175878144800663
step: 120, loss: 0.023784037679433823
step: 130, loss: 0.0036749732680618763
step: 140, loss: 0.03635983541607857
step: 150, loss: 0.09108547121286392
step: 160, loss: 0.023320410400629044
step: 170, loss: 0.1171141043305397
step: 180, loss: 0.04509983956813812
step: 190, loss: 0.03163988143205643
step: 200, loss: 0.11868558824062347
step: 210, loss: 0.027212202548980713
step: 220, loss: 0.14726832509040833
step: 230, loss: 0.09413886070251465
step: 240, loss: 0.05272320285439491
step: 250, loss: 0.04826754331588745
step: 260, loss: 0.025786811485886574
step: 270, loss: 0.007249115966260433
step: 280, loss: 0.13909709453582764
step: 290, loss: 0.007489943411201239
step: 300, loss: 0.05073947086930275
step: 310, loss: 0.09635603427886963
step: 320, loss: 0.01161795761436224
step: 330, loss: 0.03468291461467743
step: 340, loss: 0.11880150437355042
step: 350, loss: 0.006507757119834423
step: 360, loss: 0.03151267021894455
step: 370, loss: 0.021062003448605537
step: 380, loss: 0.01682894304394722
step: 390, loss: 0.09857942909002304
step: 400, loss: 0.03007015772163868
step: 410, loss: 0.020101027563214302
step: 420, loss: 0.1136523187160492
step: 430, loss: 0.022435953840613365
step: 440, loss: 0.007535920944064856
step: 450, loss: 0.017144273966550827
step: 460, loss: 0.09781556576490402
step: 470, loss: 0.22453024983406067
step: 480, loss: 0.06658809632062912
step: 490, loss: 0.056157030165195465
step: 500, loss: 0.008106104098260403

```

step: 510, loss: 0.019388431683182716
step: 520, loss: 0.07955234497785568
step: 530, loss: 0.1162550300359726
/usr/local/lib/python3.7/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=0.9066543438077633, f1=0.8960074280408542,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021881727501749992
step: 10, loss: 0.004962571896612644
step: 20, loss: 0.006401402410119772
step: 30, loss: 0.03853942081332207
step: 40, loss: 0.021822353824973106
step: 50, loss: 0.0542147159576416
step: 60, loss: 0.044324930757284164
step: 70, loss: 0.10388968884944916
step: 80, loss: 0.0117641007527709
step: 90, loss: 0.17964288592338562
step: 100, loss: 0.04451309144496918
step: 110, loss: 0.07071875780820847
step: 120, loss: 0.0027340035885572433
step: 130, loss: 0.0026846646796911955
step: 140, loss: 0.005228231195360422
step: 150, loss: 0.03734550625085831
step: 160, loss: 0.16002888977527618
step: 170, loss: 0.013710063882172108
step: 180, loss: 0.02436267025768757
step: 190, loss: 0.036349087953567505
step: 200, loss: 0.016854021698236465
step: 210, loss: 0.016384819522500038
step: 220, loss: 0.018056070432066917
step: 230, loss: 0.01815945655107498
step: 240, loss: 0.025157101452350616
step: 250, loss: 0.034887153655290604
step: 260, loss: 0.005412920843809843
step: 270, loss: 0.00662154104560614
step: 280, loss: 0.07899673283100128
step: 290, loss: 0.092631496489048
step: 300, loss: 0.002036913763731718
step: 310, loss: 0.007387756835669279
step: 320, loss: 0.04351196065545082
step: 330, loss: 0.06293956190347672
step: 340, loss: 0.03418156877160072
step: 350, loss: 0.02386840432882309
step: 360, loss: 0.1066058874130249

```

```

step: 370, loss: 0.0034725258592516184
step: 380, loss: 0.01297271903604269
step: 390, loss: 0.15821895003318787
step: 400, loss: 0.01841052435338497
step: 410, loss: 0.01743965595960617
step: 420, loss: 0.00311901792883873
step: 430, loss: 0.2066938579082489
step: 440, loss: 0.0039011952467262745
step: 450, loss: 0.0029760212637484074
step: 460, loss: 0.10907460004091263
step: 470, loss: 0.014821324497461319
step: 480, loss: 0.020936105400323868
step: 490, loss: 0.006861590780317783
step: 500, loss: 0.0045210979878902435
step: 510, loss: 0.17379331588745117
step: 520, loss: 0.07557747513055801
step: 530, loss: 0.06734079122543335
/usr/local/lib/python3.7/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=0.9038010323791646, f1=0.8959020254357042,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06234800070524216
step: 10, loss: 0.05198837071657181
step: 20, loss: 0.009055341593921185
step: 30, loss: 0.004390450660139322
step: 40, loss: 0.018661612644791603
step: 50, loss: 0.017178736627101898
step: 60, loss: 0.13876129686832428
step: 70, loss: 0.04946108162403107
step: 80, loss: 0.003979919012635946
step: 90, loss: 0.008862336166203022
step: 100, loss: 0.0007377120782621205
step: 110, loss: 0.0006393186049535871
step: 120, loss: 0.00695380661636591
step: 130, loss: 0.001359764952212572
step: 140, loss: 0.005395587533712387
step: 150, loss: 0.002214656677097082
step: 160, loss: 0.01013004407286644
step: 170, loss: 0.016901567578315735
step: 180, loss: 0.0022610535379499197
step: 190, loss: 0.003357032546773553
step: 200, loss: 0.010566947050392628
step: 210, loss: 0.09834110736846924
step: 220, loss: 0.003159309271723032

```

```

step: 230, loss: 0.01950533129274845
step: 240, loss: 0.02693227306008339
step: 250, loss: 0.00968747679144144
step: 260, loss: 0.01632056199014187
step: 270, loss: 0.057040125131607056
step: 280, loss: 0.0706232413649559
step: 290, loss: 0.12963725626468658
step: 300, loss: 0.03659576550126076
step: 310, loss: 0.0766153559088707
step: 320, loss: 0.06268524378538132
step: 330, loss: 0.022656038403511047
step: 340, loss: 0.01889481209218502
step: 350, loss: 0.015765925869345665
step: 360, loss: 0.02804413251578808
step: 370, loss: 0.04239439219236374
step: 380, loss: 0.0210331492125988
step: 390, loss: 0.020128212869167328
step: 400, loss: 0.008174861781299114
step: 410, loss: 0.052343688905239105
step: 420, loss: 0.002475901274010539
step: 430, loss: 0.019253859296441078
step: 440, loss: 0.01793147623538971
step: 450, loss: 0.11206533014774323
step: 460, loss: 0.0006226467667147517
step: 470, loss: 0.0010503209196031094
step: 480, loss: 0.024347729980945587
step: 490, loss: 0.0005402425304055214
step: 500, loss: 0.08228352665901184
step: 510, loss: 0.041682954877614975
step: 520, loss: 0.10599327087402344
step: 530, loss: 0.005637845024466515
/usr/local/lib/python3.7/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=0.9002298850574711, f1=0.8948587308939323,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007972949533723295
step: 10, loss: 0.014760933816432953
step: 20, loss: 0.06419245898723602
step: 30, loss: 0.01987736113369465
step: 40, loss: 0.03309580311179161
step: 50, loss: 0.001229586428962648
step: 60, loss: 0.0010162824764847755
step: 70, loss: 0.020677903667092323
step: 80, loss: 0.06152312457561493

```

```
step: 90, loss: 0.00202169013209641
step: 100, loss: 0.008878345601260662
step: 110, loss: 0.004423990845680237
step: 120, loss: 0.0010813501430675387
step: 130, loss: 0.042356766760349274
step: 140, loss: 0.020086314529180527
step: 150, loss: 0.0012762006372213364
step: 160, loss: 0.0008261220064014196
step: 170, loss: 0.007426067721098661
step: 180, loss: 0.09918694198131561
step: 190, loss: 0.003845100523903966
step: 200, loss: 0.006690799258649349
step: 210, loss: 0.024541044607758522
step: 220, loss: 0.04905351623892784
step: 230, loss: 0.0014237182913348079
step: 240, loss: 0.076429083943367
step: 250, loss: 0.005407672841101885
step: 260, loss: 0.002900676103308797
step: 270, loss: 0.04230000078678131
step: 280, loss: 0.016899926587939262
step: 290, loss: 0.002922364976257086
step: 300, loss: 0.0382685624063015
step: 310, loss: 0.01635373942553997
step: 320, loss: 0.00094269507098943
step: 330, loss: 0.0007973629399202764
step: 340, loss: 0.030945373699069023
step: 350, loss: 0.013318298384547234
step: 360, loss: 0.004401795100420713
step: 370, loss: 0.14290910959243774
step: 380, loss: 0.00028863479383289814
step: 390, loss: 0.009633122943341732
step: 400, loss: 0.0090206079185009
step: 410, loss: 0.0034457077272236347
step: 420, loss: 0.019732261076569557
step: 430, loss: 0.002081755083054304
step: 440, loss: 0.005959570873528719
step: 450, loss: 0.005505356937646866
step: 460, loss: 0.008239807561039925
step: 470, loss: 0.022039368748664856
step: 480, loss: 0.000743542448617518
step: 490, loss: 0.013436397537589073
step: 500, loss: 0.0023093887139111757
step: 510, loss: 0.0049505578354001045
step: 520, loss: 0.003996171057224274
step: 530, loss: 0.1359521895647049
/usr/local/lib/python3.7/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=0.9042995839112344, f1=0.903435468895079,  
best_f1=0.9001865671641792  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0037182907108217478  
step: 10, loss: 0.012729565612971783  
step: 20, loss: 0.007523941341787577  
step: 30, loss: 0.012221158482134342  
step: 40, loss: 0.12593509256839752  
step: 50, loss: 0.041698817163705826  
step: 60, loss: 0.00888262502849102  
step: 70, loss: 0.0019695607479661703  
step: 80, loss: 0.02543737180531025  
step: 90, loss: 0.004851169418543577  
step: 100, loss: 0.007883034646511078  
step: 110, loss: 0.0018699390348047018  
step: 120, loss: 0.017813790589571  
step: 130, loss: 0.011069083586335182  
step: 140, loss: 0.009177698753774166  
step: 150, loss: 0.008160402998328209  
step: 160, loss: 0.00022064118820708245  
step: 170, loss: 0.007610981352627277  
step: 180, loss: 0.0006145818624645472  
step: 190, loss: 0.0005070537445135415  
step: 200, loss: 0.00037453469121828675  
step: 210, loss: 0.002852892968803644  
step: 220, loss: 0.0034845625050365925  
step: 230, loss: 0.0027860249392688274  
step: 240, loss: 0.003824575338512659  
step: 250, loss: 0.004058892372995615  
step: 260, loss: 0.0007870797999203205  
step: 270, loss: 0.02760918438434601  
step: 280, loss: 0.0013049846747890115  
step: 290, loss: 0.0031938087195158005  
step: 300, loss: 0.0034938184544444084  
step: 310, loss: 0.00045440898975357413  
step: 320, loss: 0.0008194519905373454  
step: 330, loss: 0.010260188952088356  
step: 340, loss: 0.015529262833297253  
step: 350, loss: 0.0017310602124780416  
step: 360, loss: 0.007567902095615864  
step: 370, loss: 0.0005248228553682566  
step: 380, loss: 0.003049260936677456  
step: 390, loss: 0.00023511532344855368  
step: 400, loss: 0.00039136281702667475  
step: 410, loss: 0.003475914243608713  
step: 420, loss: 0.00267986161634326
```

```

step: 430, loss: 0.00035076745552942157
step: 440, loss: 0.003701386973261833
step: 450, loss: 0.0015118415467441082
step: 460, loss: 0.001170757575891912
step: 470, loss: 0.001630681101232767
step: 480, loss: 0.30816447734832764
step: 490, loss: 0.03320273384451866
step: 500, loss: 0.00185650575440377
step: 510, loss: 0.007479420863091946
step: 520, loss: 0.0040651182644069195
step: 530, loss: 0.007166590541601181
/usr/local/lib/python3.7/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=0.9001883239171375, f1=0.895536562203229,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008511016494594514
step: 10, loss: 0.0007651773630641401
step: 20, loss: 0.000336375436745584
step: 30, loss: 0.0013359440490603447
step: 40, loss: 0.00014967568858992308
step: 50, loss: 0.20884504914283752
step: 60, loss: 0.0005118579138070345
step: 70, loss: 0.0003033093817066401
step: 80, loss: 0.0013233220670372248
step: 90, loss: 0.000981807243078947
step: 100, loss: 0.03134314343333244
step: 110, loss: 0.000521885056514293
step: 120, loss: 0.003994151018559933
step: 130, loss: 0.0032538643572479486
step: 140, loss: 0.003124895039945841
step: 150, loss: 0.002934791147708893
step: 160, loss: 0.0063347285613417625
step: 170, loss: 0.001716942759230733
step: 180, loss: 0.0005752784782089293
step: 190, loss: 0.004232028499245644
step: 200, loss: 0.004835488740354776
step: 210, loss: 0.028054941445589066
step: 220, loss: 0.009874164126813412
step: 230, loss: 0.019860904663801193
step: 240, loss: 0.004469159059226513
step: 250, loss: 0.11076641827821732
step: 260, loss: 0.0006820371490903199
step: 270, loss: 0.0034374764654785395
step: 280, loss: 0.01667569950222969

```

```

step: 290, loss: 0.006346337962895632
step: 300, loss: 0.0012994520366191864
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 310, loss: 0.0006261547096073627
step: 320, loss: 0.1633630394935608
step: 330, loss: 0.009764902293682098
step: 340, loss: 0.00016011891420930624
step: 350, loss: 0.0015543471090495586
step: 360, loss: 0.023372551426291466
step: 370, loss: 0.0005627580685541034
step: 380, loss: 0.001493207411840558
step: 390, loss: 0.12202434986829758
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 0.018815677613019943
step: 410, loss: 0.00373090710490942
step: 420, loss: 0.0025288795586675406
step: 430, loss: 0.04852035269141197
step: 440, loss: 0.01801927387714386
step: 450, loss: 0.0003461475425865501
step: 460, loss: 0.0032709145452827215
step: 470, loss: 0.0005362137453630567
step: 480, loss: 0.004789689555764198
step: 490, loss: 0.011539501138031483
step: 500, loss: 0.0008638477884232998
step: 510, loss: 0.00011654097033897415
step: 520, loss: 0.0006416017422452569
step: 530, loss: 0.017017286270856857
/usr/local/lib/python3.7/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=0.9076233183856501, f1=0.8935788055680288,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032365107908844948
step: 10, loss: 0.00019944690575357527
step: 20, loss: 0.001759332837536931
step: 30, loss: 0.0015193312428891659
step: 40, loss: 0.012072018347680569
step: 50, loss: 0.0017829900607466698
step: 60, loss: 0.05464686080813408
step: 70, loss: 0.007356193382292986
step: 80, loss: 0.039380915462970734
step: 90, loss: 0.00011636658746283501
step: 100, loss: 0.0022259384859353304
step: 110, loss: 0.0004432521527633071
step: 120, loss: 0.016123997047543526

```

```

step: 130, loss: 0.00577665027230978
step: 140, loss: 0.003330050967633724
step: 150, loss: 0.015602074563503265
step: 160, loss: 0.0004338570579420775
step: 170, loss: 0.04258248955011368
step: 180, loss: 0.008645497262477875
step: 190, loss: 0.0028033254202455282
step: 200, loss: 0.0029638963751494884
step: 210, loss: 0.0002918118261732161
step: 220, loss: 0.10331638157367706
step: 230, loss: 0.00036752078449353576
step: 240, loss: 0.007253124378621578
step: 250, loss: 0.0010675139492377639
step: 260, loss: 0.006060771178454161
step: 270, loss: 0.0002636149001773447
step: 280, loss: 0.00016564704128541052
step: 290, loss: 0.07102254778146744
step: 300, loss: 0.00026338110910728574
step: 310, loss: 0.014186156913638115
step: 320, loss: 0.0007174713537096977
step: 330, loss: 0.0032761257607489824
step: 340, loss: 0.0005808615242131054
step: 350, loss: 0.0005431747413240373
step: 360, loss: 0.00042667650268413126
step: 370, loss: 0.0004574827034957707
step: 380, loss: 0.00037434755358844995
step: 390, loss: 0.013325809501111507
step: 400, loss: 0.0124280396848917
step: 410, loss: 0.00043853186070919037
step: 420, loss: 6.8814551923424e-05
step: 430, loss: 0.00015227538824547082
step: 440, loss: 0.0017716799629852176
step: 450, loss: 0.005371576640754938
step: 460, loss: 0.003609459614381194
step: 470, loss: 0.000173010746948421
step: 480, loss: 0.00018139970779884607
step: 490, loss: 0.011674979701638222
step: 500, loss: 0.0005632582469843328
step: 510, loss: 0.004834021907299757
step: 520, loss: 0.00023929221788421273
step: 530, loss: 0.00022783367603551596
/usr/local/lib/python3.7/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=0.9095890410958903, f1=0.9025826914363388,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005105077871121466
step: 10, loss: 5.563894228544086e-05
step: 20, loss: 0.00011706569057423621
step: 30, loss: 0.0021153653506189585
step: 40, loss: 0.01188261341303587
step: 50, loss: 0.00013607183063868433
step: 60, loss: 9.072382817976177e-05
step: 70, loss: 0.0002723304205574095
step: 80, loss: 0.002893652068451047
step: 90, loss: 0.0010397388832643628
step: 100, loss: 3.5672372177941725e-05
step: 110, loss: 0.00023837162007112056
step: 120, loss: 0.006110957823693752
step: 130, loss: 0.000170900413650088
step: 140, loss: 0.00430263951420784
step: 150, loss: 0.0009801835985854268
step: 160, loss: 0.00463111000135541
step: 170, loss: 0.0010024188086390495
step: 180, loss: 0.00040944275679066777
step: 190, loss: 0.0022566996049135923
step: 200, loss: 0.002245072042569518
step: 210, loss: 0.0019302316941320896
step: 220, loss: 0.021049275994300842
step: 230, loss: 0.004834457766264677
step: 240, loss: 0.0001318646245636046
step: 250, loss: 0.003760015359148383
step: 260, loss: 0.0038232519291341305
step: 270, loss: 0.00040252431062981486
step: 280, loss: 0.00028231702162884176
step: 290, loss: 9.507332288194448e-05
step: 300, loss: 2.7696962206391618e-05
step: 310, loss: 0.00553681468591094
step: 320, loss: 0.0023212034720927477
step: 330, loss: 0.0001531373563921079
step: 340, loss: 0.00016505480743944645
step: 350, loss: 0.00016280225827358663
step: 360, loss: 0.00011231720418436453
step: 370, loss: 0.005244504194706678
step: 380, loss: 0.07772433757781982
step: 390, loss: 0.00016143821994774044
step: 400, loss: 0.0005821958766318858
step: 410, loss: 0.0004140716337133199
step: 420, loss: 0.005041997414082289
step: 430, loss: 4.212757994537242e-05
step: 440, loss: 0.0009216234320774674
step: 450, loss: 0.00010441328777233139
step: 460, loss: 0.0007442766218446195
```

```

step: 470, loss: 0.00013011319970246404
step: 480, loss: 0.055772650986909866
step: 490, loss: 0.004985416773706675
step: 500, loss: 0.013978583738207817
step: 510, loss: 0.0015706740086898208
step: 520, loss: 0.004182510077953339
step: 530, loss: 0.00016234125359915197
/usr/local/lib/python3.7/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=0.9050486336266791, f1=0.9015398973401773,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003101142356172204
step: 10, loss: 0.0006546176737174392
step: 20, loss: 9.999765461543575e-05
step: 30, loss: 0.0016594532644376159
step: 40, loss: 0.02530290000140667
step: 50, loss: 6.641996878897771e-05
step: 60, loss: 0.005653532221913338
step: 70, loss: 0.00015149735554587096
step: 80, loss: 3.189151539118029e-05
step: 90, loss: 7.067211117828265e-05
step: 100, loss: 0.002656580414623022
step: 110, loss: 0.0015145629877224565
step: 120, loss: 0.00012895680265501142
step: 130, loss: 0.00043370318599045277
step: 140, loss: 0.0010571208549663424
step: 150, loss: 0.0005209152586758137
step: 160, loss: 0.00013811388635076582
step: 170, loss: 0.001199849764816463
step: 180, loss: 0.00014659928274340928
step: 190, loss: 0.0018779649399220943
step: 200, loss: 0.005755431950092316
step: 210, loss: 0.0007132099126465619
step: 220, loss: 0.00032544031273573637
step: 230, loss: 0.00019510285346768796
step: 240, loss: 0.0002457060618326068
step: 250, loss: 3.4767021134030074e-05
step: 260, loss: 3.539656972861849e-05
step: 270, loss: 0.0563511922955513
step: 280, loss: 0.0005267431843094528
step: 290, loss: 0.001742239692248404
step: 300, loss: 0.0007662518764846027
step: 310, loss: 0.00033361310488544405
step: 320, loss: 8.009850716916844e-05

```

```

step: 330, loss: 0.16233780980110168
step: 340, loss: 0.017560040578246117
step: 350, loss: 0.0001767689682310447
step: 360, loss: 6.439201388275251e-05
step: 370, loss: 0.00017784615920390934
step: 380, loss: 0.00021712607122026384
step: 390, loss: 0.00031933048740029335
step: 400, loss: 0.0002620200684759766
step: 410, loss: 0.000508829252794385
step: 420, loss: 0.002755658235400915
step: 430, loss: 0.00021036532416474074
step: 440, loss: 0.00013593326730187982
step: 450, loss: 0.00022937534959055483
step: 460, loss: 0.067533940076828
step: 470, loss: 8.144809544319287e-05
step: 480, loss: 0.0019730087369680405
step: 490, loss: 0.0007159641827456653
step: 500, loss: 0.004441472701728344
step: 510, loss: 0.05691832676529884
step: 520, loss: 0.00041272060479968786
step: 530, loss: 0.0014050152385607362
/usr/local/lib/python3.7/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=0.8971698113207547, f1=0.8949858088930936,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003421858709771186
step: 10, loss: 5.43322748853825e-05
step: 20, loss: 6.41515216557309e-05
step: 30, loss: 9.081723692361265e-05
step: 40, loss: 3.28447058564052e-05
step: 50, loss: 0.04212482273578644
step: 60, loss: 0.005819225683808327
step: 70, loss: 0.05464474484324455
step: 80, loss: 0.0033450149931013584
step: 90, loss: 6.631098221987486e-05
step: 100, loss: 0.002100010635331273
step: 110, loss: 0.00016973933088593185
step: 120, loss: 0.0001649553159950301
step: 130, loss: 0.0014281811891123652
step: 140, loss: 0.0006991797126829624
step: 150, loss: 0.00014267167716752738
step: 160, loss: 6.182058132253587e-05
step: 170, loss: 0.0002662704500835389
step: 180, loss: 3.427524279686622e-05

```

```

step: 190, loss: 0.00032088192529045045
step: 200, loss: 0.0018379648681730032
step: 210, loss: 0.00044037983752787113
step: 220, loss: 0.00012956385035067797
step: 230, loss: 5.8630768762668595e-05
step: 240, loss: 0.0018029877683147788
step: 250, loss: 0.0005219884915277362
step: 260, loss: 0.0005516815581358969
step: 270, loss: 0.0005563748418353498
step: 280, loss: 0.0002551450452301651
step: 290, loss: 0.0006990496185608208
step: 300, loss: 0.03498071804642677
step: 310, loss: 0.0008538489346392453
step: 320, loss: 0.00021983686019666493
step: 330, loss: 0.05572747066617012
step: 340, loss: 0.003133621532469988
step: 350, loss: 0.0005003591650165617
step: 360, loss: 0.00014396845654118806
step: 370, loss: 0.000388278451282531
step: 380, loss: 0.00449087331071496
step: 390, loss: 0.0001569433807162568
step: 400, loss: 0.0004727641644421965
step: 410, loss: 0.0021464424207806587
step: 420, loss: 0.0032314255367964506
step: 430, loss: 0.0009870154317468405
step: 440, loss: 0.005555338691920042
step: 450, loss: 0.0025741246063262224
step: 460, loss: 0.0002476507506798953
step: 470, loss: 0.09123009443283081
step: 480, loss: 0.002816811203956604
step: 490, loss: 0.0006282461690716445
step: 500, loss: 0.0001654636289458722
step: 510, loss: 0.16847100853919983
step: 520, loss: 7.419998291879892e-05
step: 530, loss: 0.00025680594262667
/usr/local/lib/python3.7/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=0.9012915129151292, f1=0.8988452655889146,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010942780412733555
step: 10, loss: 0.00011073511268477887
step: 20, loss: 0.010790865868330002
step: 30, loss: 0.00028261897386983037
step: 40, loss: 0.008628936484456062

```


step: 50, loss: 0.02447483316063881
step: 60, loss: 0.003000261727720499
step: 70, loss: 7.699218986090273e-05
step: 80, loss: 0.0012227409752085805
step: 90, loss: 0.013097535818815231
step: 100, loss: 0.00020975717052351683
step: 110, loss: 0.0009880707366392016
step: 120, loss: 0.0005740839405916631
step: 130, loss: 6.250439764698967e-05
step: 140, loss: 3.231214213883504e-05
step: 150, loss: 0.0014001975068822503
step: 160, loss: 0.00025458980235271156
step: 170, loss: 8.331897697644308e-05
step: 180, loss: 6.137338641565293e-05
step: 190, loss: 6.797032983740792e-05
step: 200, loss: 4.0182621887652203e-05
step: 210, loss: 0.000234450853895396
step: 220, loss: 6.83814796502702e-05
step: 230, loss: 8.88392241904512e-05
step: 240, loss: 8.784091915003955e-05
step: 250, loss: 5.947773752268404e-05
step: 260, loss: 0.00019786288612522185
step: 270, loss: 5.625648554996587e-05
step: 280, loss: 5.8743211411638185e-05
step: 290, loss: 3.844713864964433e-05
step: 300, loss: 0.0027712977025657892
step: 310, loss: 0.0018502651946619153
step: 320, loss: 0.0014315992593765259
step: 330, loss: 7.665391603950411e-05
step: 340, loss: 0.0032246760092675686
step: 350, loss: 4.369047383079305e-05
step: 360, loss: 3.74099072359968e-05
step: 370, loss: 0.0005471889744512737
step: 380, loss: 0.0003864279424306005
step: 390, loss: 0.000617943296674639
step: 400, loss: 0.00011397709749871865
step: 410, loss: 0.0011419332586228848
step: 420, loss: 3.271088280598633e-05
step: 430, loss: 0.00024529380607418716
step: 440, loss: 0.0004512008454184979
step: 450, loss: 0.0012543504126369953
step: 460, loss: 0.0033637480810284615
step: 470, loss: 4.770655505126342e-05
step: 480, loss: 6.537104491144419e-05
step: 490, loss: 0.00019773488747887313
step: 500, loss: 8.868631266523153e-05
step: 510, loss: 0.00021016197570133954
step: 520, loss: 0.01437081303447485

```

step: 530, loss: 0.000534422870259732
/usr/local/lib/python3.7/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=0.9008419083255379, f1=0.9037106622827618,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017449954757466912
step: 10, loss: 0.00012578791938722134
step: 20, loss: 0.0001476160396123305
step: 30, loss: 0.0011098457034677267
step: 40, loss: 0.06431430578231812
step: 50, loss: 0.0003373190702404827
step: 60, loss: 3.719900632859208e-05
step: 70, loss: 0.00014838657807558775
step: 80, loss: 0.0007865023799240589
step: 90, loss: 0.00011348428233759478
step: 100, loss: 0.027335643768310547
step: 110, loss: 4.7076573537196964e-05
step: 120, loss: 0.00015399089897982776
step: 130, loss: 0.0021892639342695475
step: 140, loss: 0.000418120704125613
step: 150, loss: 5.814810356241651e-05
step: 160, loss: 0.0004685079329647124
step: 170, loss: 0.059866707772016525
step: 180, loss: 0.00011336020543240011
step: 190, loss: 0.0025438901502639055
step: 200, loss: 0.00025321514112874866
step: 210, loss: 0.005313960835337639
step: 220, loss: 0.000352448842022568
step: 230, loss: 0.0002445682475809008
step: 240, loss: 0.0002194096305174753
step: 250, loss: 0.00018308179278392345
step: 260, loss: 0.0004804704512935132
step: 270, loss: 0.0007361077005043626
step: 280, loss: 3.8867423427291214e-05
step: 290, loss: 0.0036214643623679876
step: 300, loss: 0.0007682321011088789
step: 310, loss: 0.1155904084444046
step: 320, loss: 0.0005684178322553635
step: 330, loss: 0.108134925365448
step: 340, loss: 2.7033427613787353e-05
step: 350, loss: 0.00010808102524606511
step: 360, loss: 0.021369028836488724
step: 370, loss: 7.014313450781628e-05
step: 380, loss: 0.0005350597202777863

```

```

step: 390, loss: 0.006325055845081806
step: 400, loss: 0.00022584162070415914
step: 410, loss: 6.413146911654621e-05
step: 420, loss: 0.0011281461920589209
step: 430, loss: 0.04107372835278511
step: 440, loss: 0.03042558766901493
step: 450, loss: 0.009025519713759422
step: 460, loss: 0.003151091979816556
step: 470, loss: 0.002686128718778491
step: 480, loss: 7.973105675773695e-05
step: 490, loss: 3.7931575207039714e-05
step: 500, loss: 9.025672625284642e-05
step: 510, loss: 0.08690924197435379
step: 520, loss: 0.0002109623164869845
step: 530, loss: 6.215509347384796e-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=0.9054680259499536, f1=0.906030855539972,
best_f1=0.9001865671641792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.5551725381519645e-05
step: 10, loss: 5.11924663442187e-05
step: 20, loss: 0.0001458828483009711
step: 30, loss: 0.0001904760574689135
step: 40, loss: 0.04489728808403015
step: 50, loss: 0.00222382671199739
step: 60, loss: 0.00010589348676148802
step: 70, loss: 4.5989625505171716e-05
step: 80, loss: 0.0003422696318011731
step: 90, loss: 4.491328581934795e-05
step: 100, loss: 0.00010166219726670533
step: 110, loss: 0.01399518083781004
step: 120, loss: 0.0002736139576882124
step: 130, loss: 0.050402771681547165
step: 140, loss: 8.106456516543403e-05
step: 150, loss: 0.020074572414159775
step: 160, loss: 0.00021154488786123693
step: 170, loss: 0.0003198180056642741
step: 180, loss: 0.00026398972840979695
step: 190, loss: 0.00028670148458331823
step: 200, loss: 0.00012007168697891757
step: 210, loss: 0.0002463503915350884
step: 220, loss: 0.015564316883683205
step: 230, loss: 0.0003271533059887588
step: 240, loss: 0.021966177970170975

```

```

step: 250, loss: 0.003978559281677008
step: 260, loss: 0.00037121973582543433
step: 270, loss: 0.00040731122135184705
step: 280, loss: 0.00012399826664477587
step: 290, loss: 5.995462925056927e-05
step: 300, loss: 0.027766931802034378
step: 310, loss: 0.0003127002564724535
step: 320, loss: 0.03353894501924515
step: 330, loss: 4.400426405481994e-05
step: 340, loss: 7.888576510595158e-05
step: 350, loss: 0.002224139403551817
step: 360, loss: 0.0011509746545925736
step: 370, loss: 6.206931720953435e-05
step: 380, loss: 0.0004254210798535496
step: 390, loss: 0.026730960234999657
step: 400, loss: 0.0019082850776612759
step: 410, loss: 0.00018683211237657815
step: 420, loss: 0.00038486532866954803
step: 430, loss: 0.0009460730943828821
step: 440, loss: 0.007035884540528059
step: 450, loss: 5.760123531217687e-05
step: 460, loss: 0.0020875493064522743
step: 470, loss: 0.0003655474865809083
step: 480, loss: 3.8664162275381386e-05
step: 490, loss: 0.0020943733397871256
step: 500, loss: 0.00017200522415805608
step: 510, loss: 0.00045385034172795713
step: 520, loss: 3.207710324204527e-05
step: 530, loss: 2.6270210582879372e-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=0.9061784897025172, f1=0.9049150206706478,
best_f1=0.9001865671641792
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:32, 175.35it/s]
load_f1 = 0.9089193015573384
real_f1 = 0.9061032863849765
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 181.88it/s]

```

4.1.4 Amazon-Google - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result

in PyTorch skipping the first value of the learning rate schedule. See more

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

```

rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
  UserWarning)
step: 0, loss: 0.5603142976760864
step: 10, loss: 0.36666882038116455
step: 20, loss: 0.38503167033195496
step: 30, loss: 0.2950451672077179
step: 40, loss: 0.1882036030292511
step: 50, loss: 0.43594881892204285
step: 60, loss: 0.3095202147960663
step: 70, loss: 0.2673228979110718
step: 80, loss: 0.19683505594730377
step: 90, loss: 0.48345255851745605
step: 100, loss: 0.5528759956359863
step: 110, loss: 0.2940233647823334
step: 120, loss: 0.2764367163181305
step: 130, loss: 0.3492065966129303
step: 140, loss: 0.23523515462875366
step: 150, loss: 0.2982051372528076
step: 160, loss: 0.37269696593284607
step: 170, loss: 0.3443286716938019
step: 180, loss: 0.18641380965709686
step: 190, loss: 0.2288391888141632
step: 200, loss: 0.32214874029159546
step: 210, loss: 0.19597022235393524
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4175824175824176, 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: 0.15005554258823395
step: 10, loss: 0.3149722218513489
step: 20, loss: 0.2975628077983856
step: 30, loss: 0.3469946086406708
step: 40, loss: 0.2540143132209778
step: 50, loss: 0.16440320014953613
step: 60, loss: 0.41868430376052856
step: 70, loss: 0.132035031914711
step: 80, loss: 0.3709642291069031
step: 90, loss: 0.1883493959903717
step: 100, loss: 0.07340341806411743

```

```

step: 110, loss: 0.24218221008777618
step: 120, loss: 0.13777215778827667
step: 130, loss: 0.02989266626536846
step: 140, loss: 0.28500109910964966
step: 150, loss: 0.35375601053237915
step: 160, loss: 0.27919405698776245
step: 170, loss: 0.2021656483411789
step: 180, loss: 0.2841641306877136
step: 190, loss: 0.25904586911201477
step: 200, loss: 0.05565865337848663
step: 210, loss: 0.22819538414478302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45396825396825397, f1=0.4871794871794871,
best_f1=0.4871794871794871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12037896364927292
step: 10, loss: 0.2633518874645233
step: 20, loss: 0.30321165919303894
step: 30, loss: 0.258314311504364
step: 40, loss: 0.10814705491065979
step: 50, loss: 0.10449633002281189
step: 60, loss: 0.27178695797920227
step: 70, loss: 0.3842102885246277
step: 80, loss: 0.14447258412837982
step: 90, loss: 0.21645820140838623
step: 100, loss: 0.2453952133655548
step: 110, loss: 0.29357847571372986
step: 120, loss: 0.24320314824581146
step: 130, loss: 0.14202068746089935
step: 140, loss: 0.18406878411769867
step: 150, loss: 0.3414132297039032
step: 160, loss: 0.06808461993932724
step: 170, loss: 0.1644599735736847
step: 180, loss: 0.17851293087005615
step: 190, loss: 0.26943787932395935
step: 200, loss: 0.18440011143684387
step: 210, loss: 0.1546393781900406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```



```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4689165186500889, f1=0.5533453887884268,
best_f1=0.5533453887884268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1759122908115387
step: 10, loss: 0.20628531277179718
step: 20, loss: 0.24770981073379517
step: 30, loss: 0.13634979724884033
step: 40, loss: 0.10114428400993347
step: 50, loss: 0.1808720976114273
step: 60, loss: 0.29100295901298523
step: 70, loss: 0.2807995676994324
step: 80, loss: 0.1572612226009369
step: 90, loss: 0.05437071621417999
step: 100, loss: 0.2533729672431946
step: 110, loss: 0.26518163084983826
step: 120, loss: 0.12367476522922516
step: 130, loss: 0.11531735211610794
step: 140, loss: 0.21007320284843445
step: 150, loss: 0.09725422412157059
step: 160, loss: 0.04910008981823921
step: 170, loss: 0.2625880241394043
step: 180, loss: 0.5227002501487732
step: 190, loss: 0.2567097544670105
step: 200, loss: 0.20956899225711823
step: 210, loss: 0.20550581812858582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.48822269807280516, f1=0.52, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15224336087703705
step: 10, loss: 0.2100471407175064
step: 20, loss: 0.31190839409828186
step: 30, loss: 0.13531875610351562
step: 40, loss: 0.03542237728834152
step: 50, loss: 0.09262314438819885

```

```

step: 60, loss: 0.08322994410991669
step: 70, loss: 0.12809431552886963
step: 80, loss: 0.12935705482959747
step: 90, loss: 0.07993782311677933
step: 100, loss: 0.021113138645887375
step: 110, loss: 0.25112491846084595
step: 120, loss: 0.10266582667827606
step: 130, loss: 0.15696707367897034
step: 140, loss: 0.16787973046302795
step: 150, loss: 0.14178933203220367
step: 160, loss: 0.1428210586309433
step: 170, loss: 0.21911898255348206
step: 180, loss: 0.10956905037164688
step: 190, loss: 0.048308879137039185
step: 200, loss: 0.18872523307800293
step: 210, loss: 0.033164042979478836
/usr/local/lib/python3.7/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=0.46399999999999997, f1=0.5296950240770465, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04855899140238762
step: 10, loss: 0.11805061995983124
step: 20, loss: 0.09175628423690796
step: 30, loss: 0.06840275228023529
step: 40, loss: 0.04336833208799362
step: 50, loss: 0.007430600468069315
step: 60, loss: 0.31951919198036194
step: 70, loss: 0.01189244445413351
step: 80, loss: 0.09869330376386642
step: 90, loss: 0.21099241077899933
step: 100, loss: 0.01919412985444069
step: 110, loss: 0.13956309854984283
step: 120, loss: 0.06998101621866226
step: 130, loss: 0.14667953550815582
step: 140, loss: 0.09124105423688889
step: 150, loss: 0.034398917108774185
step: 160, loss: 0.022212523967027664
step: 170, loss: 0.25189194083213806
step: 180, loss: 0.16062703728675842
step: 190, loss: 0.10622561722993851
step: 200, loss: 0.02558109723031521
step: 210, loss: 0.07814373821020126
/usr/local/lib/python3.7/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=0.4745166959578207, f1=0.5065666041275797, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010611608624458313
step: 10, loss: 0.08622027933597565
step: 20, loss: 0.0312496330589056
step: 30, loss: 0.008778839372098446
step: 40, loss: 0.05726603418588638
step: 50, loss: 0.17291942238807678
step: 60, loss: 0.08082418143749237
step: 70, loss: 0.0569775253534317
step: 80, loss: 0.02636174112558365
step: 90, loss: 0.044230084866285324
step: 100, loss: 0.00945740845054388
step: 110, loss: 0.24227473139762878
step: 120, loss: 0.16955187916755676
step: 130, loss: 0.027337871491909027
step: 140, loss: 0.00849417969584465
step: 150, loss: 0.057410866022109985
step: 160, loss: 0.08801393955945969
step: 170, loss: 0.048876870423555374
step: 180, loss: 0.03823940455913544
step: 190, loss: 0.16206659376621246
step: 200, loss: 0.03679549694061279
step: 210, loss: 0.10709558427333832
/usr/local/lib/python3.7/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=0.45960502692998206, f1=0.5065176908752328, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052634332329034805
step: 10, loss: 0.03825884684920311
step: 20, loss: 0.03217284381389618
step: 30, loss: 0.01908666267991066
step: 40, loss: 0.13267840445041656
step: 50, loss: 0.025864819064736366
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.11006324738264084
step: 70, loss: 0.13361303508281708
step: 80, loss: 0.07967530190944672
step: 90, loss: 0.03274637088179588
step: 100, loss: 0.11509286612272263
step: 110, loss: 0.11825942993164062
step: 120, loss: 0.010085396468639374
step: 130, loss: 0.0066104610450565815

```

```

step: 140, loss: 0.016952181234955788
step: 150, loss: 0.15196512639522552
step: 160, loss: 0.07721802592277527
step: 170, loss: 0.07482615113258362
step: 180, loss: 0.0669717937707901
step: 190, loss: 0.014927428215742111
step: 200, loss: 0.012682772241532803
step: 210, loss: 0.1549387276172638
/usr/local/lib/python3.7/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=0.45901639344262296, f1=0.49019607843137253, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06883305311203003
step: 10, loss: 0.09131260961294174
step: 20, loss: 0.013441899791359901
step: 30, loss: 0.03663768246769905
step: 40, loss: 0.039064664393663406
step: 50, loss: 0.041956428438425064
step: 60, loss: 0.11185549944639206
step: 70, loss: 0.009658575057983398
step: 80, loss: 0.005921050440520048
step: 90, loss: 0.02183365635573864
step: 100, loss: 0.11215405911207199
step: 110, loss: 0.09230181574821472
step: 120, loss: 0.036027777940034866
step: 130, loss: 0.053908493369817734
step: 140, loss: 0.10002695024013519
step: 150, loss: 0.22642090916633606
step: 160, loss: 0.014289802871644497
step: 170, loss: 0.03429523855447769
step: 180, loss: 0.006571552250534296
step: 190, loss: 0.003235441166907549
step: 200, loss: 0.042923279106616974
step: 210, loss: 0.010062080807983875
/usr/local/lib/python3.7/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=0.4338235294117647, f1=0.48275862068965514, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040368176996707916
step: 10, loss: 0.013260009698569775
step: 20, loss: 0.003030425636097789
step: 30, loss: 0.15991009771823883

```

```
step: 40, loss: 0.006496886257082224
step: 50, loss: 0.0256675835698843
step: 60, loss: 0.1452576220035553
step: 70, loss: 0.10945295542478561
step: 80, loss: 0.056387919932603836
step: 90, loss: 0.05892856419086456
step: 100, loss: 0.012852579355239868
step: 110, loss: 0.06753519177436829
step: 120, loss: 0.11205055564641953
step: 130, loss: 0.016216447576880455
step: 140, loss: 0.035694532096385956
step: 150, loss: 0.10696498304605484
step: 160, loss: 0.02652730792760849
step: 170, loss: 0.03957125172019005
step: 180, loss: 0.09493769705295563
step: 190, loss: 0.1740482598543167
step: 200, loss: 0.0018003208097070456
step: 210, loss: 0.11529024690389633
/usr/local/lib/python3.7/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=0.4521072796934866, f1=0.46034816247582205, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00627094553783536
step: 10, loss: 0.09059964120388031
step: 20, loss: 0.04808780178427696
step: 30, loss: 0.007028873544186354
step: 40, loss: 0.008630679920315742
step: 50, loss: 0.0035396211314946413
step: 60, loss: 0.003811864648014307
step: 70, loss: 0.016624849289655685
step: 80, loss: 0.015919536352157593
step: 90, loss: 0.005448623560369015
step: 100, loss: 0.006698727607727051
step: 110, loss: 0.03467273339629173
step: 120, loss: 0.023654883727431297
step: 130, loss: 0.0032660518772900105
step: 140, loss: 0.021580561995506287
step: 150, loss: 0.0007295591058209538
step: 160, loss: 0.011786136776208878
step: 170, loss: 0.002104173181578517
step: 180, loss: 0.005981935653835535
step: 190, loss: 0.0014771174173802137
step: 200, loss: 0.004251736681908369
step: 210, loss: 0.0010810426902025938
/usr/local/lib/python3.7/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=0.4671532846715329, f1=0.4789762340036563, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0415378212928772
step: 10, loss: 0.003254712326452136
step: 20, loss: 0.003959663677960634
step: 30, loss: 0.004622709937393665
step: 40, loss: 0.05495740473270416
step: 50, loss: 0.0029369075782597065
step: 60, loss: 0.22913233935832977
step: 70, loss: 0.00041856139432638884
step: 80, loss: 0.04698508232831955
step: 90, loss: 0.026639223098754883
step: 100, loss: 0.04777710139751434
step: 110, loss: 0.004396986681967974
step: 120, loss: 0.05010238662362099
step: 130, loss: 0.0019125055987387896
step: 140, loss: 0.002703011268749833
step: 150, loss: 0.003968921024352312
step: 160, loss: 0.006808183155953884
step: 170, loss: 0.0029495935887098312
step: 180, loss: 0.006117160432040691
step: 190, loss: 0.10983137041330338
step: 200, loss: 0.00923120230436325
step: 210, loss: 0.004556534346193075
/usr/local/lib/python3.7/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=0.4394250513347023, f1=0.4551148225469729, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027730893343687057
step: 10, loss: 0.0004826000367756933
step: 20, loss: 0.004234304651618004
step: 30, loss: 0.022421278059482574
step: 40, loss: 0.011976909823715687
step: 50, loss: 0.0011160324793308973
step: 60, loss: 0.0022672698833048344
step: 70, loss: 0.020890282467007637
step: 80, loss: 0.05663115903735161
step: 90, loss: 0.0008259625174105167
step: 100, loss: 0.054300207644701004
step: 110, loss: 0.012136614881455898
step: 120, loss: 0.00859695952385664

```

```

step: 130, loss: 0.0029208653140813112
step: 140, loss: 0.03630983829498291
step: 150, loss: 0.003556040581315756
step: 160, loss: 0.0035780721809715033
step: 170, loss: 0.08355532586574554
step: 180, loss: 0.015361586585640907
step: 190, loss: 0.008698448538780212
step: 200, loss: 0.11452627927064896
step: 210, loss: 0.005531803239136934
/usr/local/lib/python3.7/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=0.45833333333333337, f1=0.4448669201520913, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010976763442158699
step: 10, loss: 0.01317105907946825
step: 20, loss: 0.005117708817124367
step: 30, loss: 0.0007260615238919854
step: 40, loss: 0.03337622061371803
step: 50, loss: 0.0023526481818407774
step: 60, loss: 0.0076688374392688274
step: 70, loss: 0.004217250272631645
step: 80, loss: 0.002240107161924243
step: 90, loss: 0.061534393578767776
step: 100, loss: 0.007149806246161461
step: 110, loss: 0.002641316968947649
step: 120, loss: 0.0033065860625356436
step: 130, loss: 0.004037213046103716
step: 140, loss: 0.02834009937942028
step: 150, loss: 0.008226846344769001
step: 160, loss: 0.11648331582546234
step: 170, loss: 0.0010018819011747837
step: 180, loss: 0.06485776603221893
step: 190, loss: 0.001863346784375608
step: 200, loss: 0.08540981262922287
step: 210, loss: 0.03362058848142624
/usr/local/lib/python3.7/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=0.44313725490196076, f1=0.45940594059405937, best_f1=0.52
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009309621527791023
step: 10, loss: 0.0017946595326066017
step: 20, loss: 0.055928051471710205

```

```

step: 30, loss: 0.05367287993431091
step: 40, loss: 0.007627400103956461
step: 50, loss: 0.0031206991989165545
step: 60, loss: 0.054228488355875015
step: 70, loss: 0.0031291248742491007
step: 80, loss: 0.0053047118708491325
step: 90, loss: 0.01986517384648323
step: 100, loss: 0.001967281335964799
step: 110, loss: 0.0013066749088466167
step: 120, loss: 0.015611999668180943
step: 130, loss: 0.010408393107354641
step: 140, loss: 0.0014922904083505273
step: 150, loss: 0.003149823285639286
step: 160, loss: 0.0022829407826066017
step: 170, loss: 0.0005138495471328497
step: 180, loss: 0.0006991632981225848
step: 190, loss: 0.01590416021645069
step: 200, loss: 0.051853202283382416
step: 210, loss: 0.018396664410829544
/usr/local/lib/python3.7/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=0.4474708171206226, f1=0.45940594059405937, best_f1=0.52
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2293it [00:09, 241.71it/s]
load_f1 = 0.48375451263537905
real_f1 = 0.48363636363636364
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 176.05it/s]

```

4.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

```
step: 0, loss: 0.5476108193397522  
step: 10, loss: 0.3836674392223358  
step: 20, loss: 0.3001682758331299  
step: 30, loss: 0.40133601427078247  
step: 40, loss: 0.4373561441898346  
step: 50, loss: 0.3090874254703522  
step: 60, loss: 0.28447261452674866  
step: 70, loss: 0.2916218340396881
```

```

step: 80, loss: 0.23680011928081512
step: 90, loss: 0.29648295044898987
step: 100, loss: 0.3198568820953369
step: 110, loss: 0.3691559135913849
step: 120, loss: 0.1119728535413742
step: 130, loss: 0.18179108202457428
step: 140, loss: 0.0702705830335617
step: 150, loss: 0.18741412460803986
step: 160, loss: 0.15221019089221954
step: 170, loss: 0.24256330728530884
step: 180, loss: 0.04553528502583504
step: 190, loss: 0.2332042157649994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6601941747572816, f1=0.676470588235294,
best_f1=0.676470588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20300674438476562
step: 10, loss: 0.11758013814687729
step: 20, loss: 0.2606046795845032
step: 30, loss: 0.16066379845142365
step: 40, loss: 0.13403992354869843
step: 50, loss: 0.03786816820502281
step: 60, loss: 0.18465930223464966
step: 70, loss: 0.33036261796951294
step: 80, loss: 0.11250382661819458
step: 90, loss: 0.12546144425868988
step: 100, loss: 0.010572666302323341
step: 110, loss: 0.25642988085746765
step: 120, loss: 0.23013979196548462
step: 130, loss: 0.1875699758529663
step: 140, loss: 0.1399373710155487
step: 150, loss: 0.17611832916736603
step: 160, loss: 0.1582515835762024
step: 170, loss: 0.17331410944461823
step: 180, loss: 0.10871793329715729
step: 190, loss: 0.04819580912590027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.7532467532467532, f1=0.7604166666666666,
best_f1=0.7604166666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16419516503810883
step: 10, loss: 0.16401232779026031
step: 20, loss: 0.06975893676280975
step: 30, loss: 0.1209322139620781
step: 40, loss: 0.1329733431339264
step: 50, loss: 0.2773014307022095
step: 60, loss: 0.07089672982692719
step: 70, loss: 0.10307082533836365
step: 80, loss: 0.05755574256181717
step: 90, loss: 0.14585727453231812
step: 100, loss: 0.11578354239463806
step: 110, loss: 0.1090267226099968
step: 120, loss: 0.08404242247343063
step: 130, loss: 0.023577041923999786
step: 140, loss: 0.019421571865677834
step: 150, loss: 0.11229224503040314
step: 160, loss: 0.012231813743710518
step: 170, loss: 0.23007714748382568
step: 180, loss: 0.029985293745994568
step: 190, loss: 0.07583452761173248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7814207650273224, f1=0.7733333333333333,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04377760365605354
step: 10, loss: 0.2085021734237671
step: 20, loss: 0.15145501494407654
step: 30, loss: 0.08476202189922333
step: 40, loss: 0.14070159196853638
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.05450260266661644
step: 60, loss: 0.18195828795433044

```

```

step: 70, loss: 0.04446347802877426
step: 80, loss: 0.22734051942825317
step: 90, loss: 0.0748639702796936
step: 100, loss: 0.12799988687038422
step: 110, loss: 0.009716206230223179
step: 120, loss: 0.054329708218574524
step: 130, loss: 0.12533603608608246
step: 140, loss: 0.0485055074095726
step: 150, loss: 0.049716148525476456
step: 160, loss: 0.004764493554830551
step: 170, loss: 0.1873626559972763
step: 180, loss: 0.08002225309610367
step: 190, loss: 0.08023517578840256
/usr/local/lib/python3.7/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=0.7794871794871795, f1=0.7487179487179487,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05244222655892372
step: 10, loss: 0.006302101071923971
step: 20, loss: 0.06946009397506714
step: 30, loss: 0.020600855350494385
step: 40, loss: 0.07532258331775665
step: 50, loss: 0.05168318748474121
step: 60, loss: 0.0631629154086113
step: 70, loss: 0.019197402521967888
step: 80, loss: 0.003217745805159211
step: 90, loss: 0.011946937069296837
step: 100, loss: 0.0024607954546809196
step: 110, loss: 0.0048562902957201
step: 120, loss: 0.018508372828364372
step: 130, loss: 0.04069555550813675
step: 140, loss: 0.03775991499423981
step: 150, loss: 0.11197112500667572
step: 160, loss: 0.02218516729772091
step: 170, loss: 0.004090507049113512
step: 180, loss: 0.03712107241153717
step: 190, loss: 0.10369753837585449
/usr/local/lib/python3.7/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=0.7580645161290323, f1=0.724233983286908,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005001209210604429
step: 10, loss: 0.008195621892809868
step: 20, loss: 0.002566884970292449
step: 30, loss: 0.0011932130437344313
step: 40, loss: 0.08367675542831421
step: 50, loss: 0.007955270819365978
step: 60, loss: 0.0008331738645210862
step: 70, loss: 0.021416518837213516
step: 80, loss: 0.12819284200668335
step: 90, loss: 0.036245640367269516
step: 100, loss: 0.00047504555550403893
step: 110, loss: 0.025891665369272232
step: 120, loss: 0.1206723153591156
step: 130, loss: 0.0043769218027591705
step: 140, loss: 0.011760788969695568
step: 150, loss: 0.050005920231342316
step: 160, loss: 0.0772458016872406
step: 170, loss: 0.04128820821642876
step: 180, loss: 0.025801455602049828
step: 190, loss: 0.004128210712224245
/usr/local/lib/python3.7/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=0.7479224376731302, f1=0.7094972067039106,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010088887996971607
step: 10, loss: 0.015290606766939163
step: 20, loss: 0.04917389154434204
step: 30, loss: 0.029755182564258575
step: 40, loss: 0.002004713285714388
step: 50, loss: 0.05594419687986374
step: 60, loss: 0.02261979877948761
step: 70, loss: 0.0013580848462879658
step: 80, loss: 0.11422105878591537
step: 90, loss: 0.001844294834882021
step: 100, loss: 0.0003935388522222638
step: 110, loss: 0.024563593789935112
step: 120, loss: 0.006569486577063799
step: 130, loss: 0.07802609354257584
step: 140, loss: 0.024180583655834198
step: 150, loss: 0.03248148784041405
step: 160, loss: 0.3143850862979889
step: 170, loss: 0.03907296061515808
step: 180, loss: 0.026105228811502457

```

```

step: 190, loss: 0.009364535100758076
/usr/local/lib/python3.7/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=0.7513513513513514, f1=0.747191011235955,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015821419656276703
step: 10, loss: 0.03463771566748619
step: 20, loss: 0.11432412266731262
step: 30, loss: 0.010424422100186348
step: 40, loss: 0.0008687812951393425
step: 50, loss: 0.0012628216063603759
step: 60, loss: 0.001089634490199387
step: 70, loss: 0.007854178547859192
step: 80, loss: 0.0013471698621287942
step: 90, loss: 0.07500284910202026
step: 100, loss: 0.02026766538619995
step: 110, loss: 0.0005845400155521929
step: 120, loss: 0.004496371373534203
step: 130, loss: 0.0029098966624587774
step: 140, loss: 0.06851150840520859
step: 150, loss: 0.009326175786554813
step: 160, loss: 0.0010405994253233075
step: 170, loss: 0.003186864545568824
step: 180, loss: 0.0030412091873586178
step: 190, loss: 0.010820893570780754
/usr/local/lib/python3.7/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=0.7556675062972291, f1=0.7512953367875648,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009421895956620574
step: 10, loss: 0.02320513315498829
step: 20, loss: 0.06611136347055435
step: 30, loss: 0.002143907593563199
step: 40, loss: 0.00022837132564745843
step: 50, loss: 0.11427812278270721
step: 60, loss: 0.0004939265782013535
step: 70, loss: 0.000693333859089762
step: 80, loss: 0.0007930430001579225
step: 90, loss: 0.005953256506472826
step: 100, loss: 0.01577134244143963

```

```

step: 110, loss: 0.025030061602592468
step: 120, loss: 0.06828510016202927
step: 130, loss: 0.0027623118367046118
step: 140, loss: 0.014109320007264614
step: 150, loss: 0.010266635566949844
step: 160, loss: 0.0020606659818440676
step: 170, loss: 0.00042318477062508464
step: 180, loss: 0.1841755211353302
step: 190, loss: 0.0059188553132116795
/usr/local/lib/python3.7/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=0.7708333333333334, f1=0.7405405405405405,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020993612706661224
step: 10, loss: 0.12149467319250107
step: 20, loss: 0.000797910732217133
step: 30, loss: 0.0015981529140844941
step: 40, loss: 0.004744702950119972
step: 50, loss: 0.004650600720196962
step: 60, loss: 0.00035837452742271125
step: 70, loss: 0.0006562244379892945
step: 80, loss: 0.0076970504596829414
step: 90, loss: 0.0018120176391676068
step: 100, loss: 0.014616432599723339
step: 110, loss: 0.001116889645345509
step: 120, loss: 0.014291157945990562
step: 130, loss: 0.0011876170756295323
step: 140, loss: 0.04453081637620926
step: 150, loss: 0.017326144501566887
step: 160, loss: 0.001827784115448594
step: 170, loss: 0.0013394642155617476
step: 180, loss: 0.0011355889728292823
step: 190, loss: 0.002398004289716482
/usr/local/lib/python3.7/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=0.7553191489361701, f1=0.7547169811320755,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004109145607799292
step: 10, loss: 0.0028224969282746315
step: 20, loss: 0.00222354126162827

```



```

step: 30, loss: 0.010742150247097015
step: 40, loss: 0.0014451813185587525
step: 50, loss: 0.06474462896585464
step: 60, loss: 0.001863220240920782
step: 70, loss: 0.0008168938802555203
step: 80, loss: 0.0009600811172276735
step: 90, loss: 0.001139993080869317
step: 100, loss: 0.0004675397649407387
step: 110, loss: 0.001239714794792235
step: 120, loss: 0.001596861518919468
step: 130, loss: 0.007260431069880724
step: 140, loss: 0.0008446322754025459
step: 150, loss: 0.0003531461989041418
step: 160, loss: 0.0009096928406506777
step: 170, loss: 0.0023797890171408653
step: 180, loss: 0.0011529254261404276
step: 190, loss: 0.0017287717200815678
/usr/local/lib/python3.7/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=0.7611548556430445, f1=0.753315649867374,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032473832834511995
step: 10, loss: 0.0011820318177342415
step: 20, loss: 0.0009519311133772135
step: 30, loss: 0.00150700518861413
step: 40, loss: 0.027057308703660965
step: 50, loss: 0.0036110880319029093
step: 60, loss: 0.0010784954065456986
step: 70, loss: 0.004510669969022274
step: 80, loss: 0.0011713167186826468
step: 90, loss: 0.0021921349689364433
step: 100, loss: 0.0008794492459855974
step: 110, loss: 0.02341325767338276
step: 120, loss: 0.0007297990960069001
step: 130, loss: 0.01702912151813507
step: 140, loss: 0.10742849856615067
step: 150, loss: 0.0008280125912278891
step: 160, loss: 0.000409287546062842
step: 170, loss: 0.0008697360171936452
step: 180, loss: 0.001482417923398316
step: 190, loss: 0.0018127256771549582
/usr/local/lib/python3.7/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=0.7613636363636365, f1=0.7485380116959065,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015712716849520802
step: 10, loss: 0.005291924811899662
step: 20, loss: 0.003906567115336657
step: 30, loss: 0.00786955002695322
step: 40, loss: 0.0007526496774517
step: 50, loss: 0.026172934100031853
step: 60, loss: 0.0022299890406429768
step: 70, loss: 0.003919829614460468
step: 80, loss: 0.0022891336120665073
step: 90, loss: 0.004459437448531389
step: 100, loss: 0.0004928825655952096
step: 110, loss: 0.0006648823618888855
step: 120, loss: 0.0005080448463559151
step: 130, loss: 0.0003290420863777399
step: 140, loss: 0.0002829233999364078
step: 150, loss: 0.00042032450437545776
step: 160, loss: 0.001810411922633648
step: 170, loss: 0.010923103429377079
step: 180, loss: 0.0005933705833740532
step: 190, loss: 0.0022968638222664595
/usr/local/lib/python3.7/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=0.7660668380462724, f1=0.7591623036649214,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018794931238517165
step: 10, loss: 0.0018351763719692826
step: 20, loss: 0.0017049201996996999
step: 30, loss: 0.0004231571510899812
step: 40, loss: 0.0006246128468774259
step: 50, loss: 0.0008170484798029065
step: 60, loss: 0.00033684720983728766
step: 70, loss: 0.000305190565995872
step: 80, loss: 0.0007633411441929638
step: 90, loss: 0.0004981064703315496
step: 100, loss: 0.0035376872401684523
step: 110, loss: 0.0005345772369764745
step: 120, loss: 0.0011341789504513144
step: 130, loss: 0.001996226143091917
step: 140, loss: 0.0009290472371503711

```

```

step: 150, loss: 0.001075263018719852
step: 160, loss: 0.007548003923147917
step: 170, loss: 0.18062669038772583
step: 180, loss: 0.0012853057123720646
step: 190, loss: 0.000696102564688772
/usr/local/lib/python3.7/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=0.7657430730478589, f1=0.7544303797468355,
best_f1=0.7733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0216436218470335
step: 10, loss: 0.019080236554145813
step: 20, loss: 0.032796528190374374
step: 30, loss: 0.003912733867764473
step: 40, loss: 0.00032225658651441336
step: 50, loss: 0.00041459081694483757
step: 60, loss: 0.033542752265930176
step: 70, loss: 0.01633651740849018
step: 80, loss: 0.01926889270544052
step: 90, loss: 0.0026547221932560205
step: 100, loss: 0.0006079343729652464
step: 110, loss: 0.003648376790806651
step: 120, loss: 0.0018224997911602259
step: 130, loss: 0.0003779701073653996
step: 140, loss: 0.002320900559425354
step: 150, loss: 0.0012834996450692415
step: 160, loss: 0.00039714283775538206
step: 170, loss: 0.0006548103410750628
step: 180, loss: 0.0005369610153138638
step: 190, loss: 0.003050813917070627
/usr/local/lib/python3.7/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=0.7700258397932817, f1=0.7539267015706806,
best_f1=0.7733333333333333
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True

```

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 160.19it/s]
load_f1 = 0.7214854111405835
real_f1 = 0.7015706806282722
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 176.83it/s]

```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6165744662284851
step: 10, loss: 0.3687102496623993
step: 20, loss: 0.30875706672668457
step: 30, loss: 0.3907177150249481
step: 40, loss: 0.28262069821357727
step: 50, loss: 0.25831300020217896
step: 60, loss: 0.24509304761886597
step: 70, loss: 0.3771314322948456
step: 80, loss: 0.3846713602542877
step: 90, loss: 0.26410335302352905
step: 100, loss: 0.2664520740509033
step: 110, loss: 0.25776904821395874
step: 120, loss: 0.09186646342277527
step: 130, loss: 0.06763725727796555
step: 140, loss: 0.25295522809028625
step: 150, loss: 0.35514819622039795
step: 160, loss: 0.164027139544487
step: 170, loss: 0.2913610339164734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6652267818574514, f1=0.6098081023454158,
best_f1=0.6098081023454158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15148545801639557
step: 10, loss: 0.13722816109657288
step: 20, loss: 0.05376318842172623
step: 30, loss: 0.20245185494422913
step: 40, loss: 0.11500225216150284
step: 50, loss: 0.12971222400665283
step: 60, loss: 0.3542972803115845
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.08057165145874023

```

```

step: 80, loss: 0.13726577162742615
step: 90, loss: 0.11182127147912979
step: 100, loss: 0.19807040691375732
step: 110, loss: 0.10835840553045273
step: 120, loss: 0.10697413235902786
step: 130, loss: 0.06557872891426086
step: 140, loss: 0.2827779948711395
step: 150, loss: 0.1896727979183197
step: 160, loss: 0.13454435765743256
step: 170, loss: 0.09140650928020477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7182044887780548, f1=0.6898263027295285,
best_f1=0.6898263027295285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1318625807762146
step: 10, loss: 0.19465714693069458
step: 20, loss: 0.024529553949832916
step: 30, loss: 0.07818819582462311
step: 40, loss: 0.11272766441106796
step: 50, loss: 0.11966900527477264
step: 60, loss: 0.25148993730545044
step: 70, loss: 0.2067977339029312
step: 80, loss: 0.09210427105426788
step: 90, loss: 0.13573536276817322
step: 100, loss: 0.008174563758075237
step: 110, loss: 0.20702724158763885
step: 120, loss: 0.0815286636352539
step: 130, loss: 0.1609739363193512
step: 140, loss: 0.10735742747783661
step: 150, loss: 0.026757989078760147
step: 160, loss: 0.09709354490041733
step: 170, loss: 0.0693373903632164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7354497354497355, f1=0.731829573934837,

```

```

best_f1=0.731829573934837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007734957616776228
step: 10, loss: 0.12581120431423187
step: 20, loss: 0.1781364381313324
step: 30, loss: 0.11441954970359802
step: 40, loss: 0.010684964247047901
step: 50, loss: 0.2994500994682312
step: 60, loss: 0.14808911085128784
step: 70, loss: 0.11398173123598099
step: 80, loss: 0.08794277906417847
step: 90, loss: 0.09174685180187225
step: 100, loss: 0.28627994656562805
step: 110, loss: 0.09272298961877823
step: 120, loss: 0.04605848714709282
step: 130, loss: 0.18176348507404327
step: 140, loss: 0.09850011765956879
step: 150, loss: 0.17477761209011078
step: 160, loss: 0.08763975650072098
step: 170, loss: 0.012660828419029713
/usr/local/lib/python3.7/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=0.7342465753424658, f1=0.7183462532299741,
best_f1=0.731829573934837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07286972552537918
step: 10, loss: 0.014609753154218197
step: 20, loss: 0.04111155495047569
step: 30, loss: 0.14772649109363556
step: 40, loss: 0.06634248793125153
step: 50, loss: 0.24812881648540497
step: 60, loss: 0.06567411124706268
step: 70, loss: 0.45413103699684143
step: 80, loss: 0.1427159607410431
step: 90, loss: 0.1188163235783577
step: 100, loss: 0.06915626674890518
step: 110, loss: 0.13011394441127777
step: 120, loss: 0.022548966109752655
step: 130, loss: 0.13299289345741272
step: 140, loss: 0.04807819798588753
step: 150, loss: 0.0728679746389389
step: 160, loss: 0.08566264063119888
step: 170, loss: 0.21110357344150543

```



```

/usr/local/lib/python3.7/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=0.7095115681233933, f1=0.712871287128713,
best_f1=0.731829573934837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006224621552973986
step: 10, loss: 0.04980183392763138
step: 20, loss: 0.005832298658788204
step: 30, loss: 0.05749956890940666
step: 40, loss: 0.00743865454569459
step: 50, loss: 0.18002742528915405
step: 60, loss: 0.04559442028403282
step: 70, loss: 0.11970579624176025
step: 80, loss: 0.12196437269449234
step: 90, loss: 0.1497727483510971
step: 100, loss: 0.0942973792552948
step: 110, loss: 0.003443132620304823
step: 120, loss: 0.04664458706974983
step: 130, loss: 0.04041968286037445
step: 140, loss: 0.021839605644345284
step: 150, loss: 0.02791237086057663
step: 160, loss: 0.11328544467687607
step: 170, loss: 0.023008624091744423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7440633245382585, f1=0.6812652068126521,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003319420153275132
step: 10, loss: 0.015664134174585342
step: 20, loss: 0.001913935411721468
step: 30, loss: 0.010261589661240578
step: 40, loss: 0.02140977792441845
step: 50, loss: 0.03734055534005165
step: 60, loss: 0.055504582822322845
step: 70, loss: 0.01582157239317894
step: 80, loss: 0.13666388392448425
step: 90, loss: 0.005818639416247606

```

```

step: 100, loss: 0.061955131590366364
step: 110, loss: 0.11575474590063095
step: 120, loss: 0.03437570482492447
step: 130, loss: 0.1401916742324829
step: 140, loss: 0.023056760430336
step: 150, loss: 0.027978813275694847
step: 160, loss: 0.03312402591109276
step: 170, loss: 0.09159255027770996
/usr/local/lib/python3.7/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=0.7219251336898396, f1=0.6911764705882352,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018802262842655182
step: 10, loss: 0.007635799236595631
step: 20, loss: 0.0010000471957027912
step: 30, loss: 0.03664612025022507
step: 40, loss: 0.004204050172120333
step: 50, loss: 0.017261000350117683
step: 60, loss: 0.019608039408922195
step: 70, loss: 0.030800271779298782
step: 80, loss: 0.013370418921113014
step: 90, loss: 0.04980875551700592
step: 100, loss: 0.07915966957807541
step: 110, loss: 0.04657420143485069
step: 120, loss: 0.0016619381494820118
step: 130, loss: 0.009432741440832615
step: 140, loss: 0.013161428272724152
step: 150, loss: 0.05808284133672714
step: 160, loss: 0.006525831762701273
step: 170, loss: 0.04723585024476051
/usr/local/lib/python3.7/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=0.7146666666666667, f1=0.6865671641791046,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0710742250084877
step: 10, loss: 0.017867859452962875
step: 20, loss: 0.007701158989220858
step: 30, loss: 0.03673524409532547
step: 40, loss: 0.0024990695528686047
step: 50, loss: 0.0012678081402555108

```

```

step: 60, loss: 0.004003297071903944
step: 70, loss: 0.07424932718276978
step: 80, loss: 0.02390962466597557
step: 90, loss: 0.04906357079744339
step: 100, loss: 0.001448155613616109
step: 110, loss: 0.001595428679138422
step: 120, loss: 0.0013137178029865026
step: 130, loss: 0.020175063982605934
step: 140, loss: 0.00869427714496851
step: 150, loss: 0.0731622725725174
step: 160, loss: 0.07043398171663284
step: 170, loss: 0.026775088161230087
/usr/local/lib/python3.7/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=0.6963788300835654, f1=0.6649484536082474,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0842069536447525
step: 10, loss: 0.0010173845803365111
step: 20, loss: 0.03575995936989784
step: 30, loss: 0.005582205951213837
step: 40, loss: 0.0008652857504785061
step: 50, loss: 0.04664987325668335
step: 60, loss: 0.002319718012586236
step: 70, loss: 0.0027914054226130247
step: 80, loss: 0.006558613386005163
step: 90, loss: 0.0014594817766919732
step: 100, loss: 0.0003113462298642844
step: 110, loss: 0.007500129286199808
step: 120, loss: 0.034442488104104996
step: 130, loss: 0.05558916553854942
step: 140, loss: 0.03560260310769081
step: 150, loss: 0.010408308357000351
step: 160, loss: 0.038295019418001175
step: 170, loss: 0.008549480699002743
/usr/local/lib/python3.7/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=0.7142857142857143, f1=0.6960784313725491,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11123666167259216
step: 10, loss: 0.020825659856200218

```

```

step: 20, loss: 0.02539539523422718
step: 30, loss: 0.0006682202802039683
step: 40, loss: 0.058161456137895584
step: 50, loss: 0.008157548494637012
step: 60, loss: 0.010536868125200272
step: 70, loss: 0.0083592738956213
step: 80, loss: 0.0015898304991424084
step: 90, loss: 0.002218155423179269
step: 100, loss: 0.009419032372534275
step: 110, loss: 0.006519257556647062
step: 120, loss: 0.0037038701120764017
step: 130, loss: 0.0052422937005758286
step: 140, loss: 0.0005412446917034686
step: 150, loss: 0.08565006405115128
step: 160, loss: 0.0051059904508292675
step: 170, loss: 0.01780877821147442
/usr/local/lib/python3.7/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=0.6927374301675977, f1=0.6700507614213198,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0045167249627411366
step: 10, loss: 0.001876979717053473
step: 20, loss: 0.005741169676184654
step: 30, loss: 0.007355926092714071
step: 40, loss: 0.006106162443757057
step: 50, loss: 0.025686083361506462
step: 60, loss: 0.054832424968481064
step: 70, loss: 0.013594276271760464
step: 80, loss: 0.0009269932634197176
step: 90, loss: 0.01857070066034794
step: 100, loss: 0.006994871888309717
step: 110, loss: 0.017931032925844193
step: 120, loss: 0.0074211303144693375
step: 130, loss: 0.014628888107836246
step: 140, loss: 0.061035312712192535
step: 150, loss: 0.1947575956583023
step: 160, loss: 0.005065571051090956
step: 170, loss: 0.0005206321948207915
/usr/local/lib/python3.7/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=0.6943765281173595, f1=0.6915887850467289,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015606630593538284
step: 10, loss: 0.11610884964466095
step: 20, loss: 0.004887885879725218
step: 30, loss: 0.004722590558230877
step: 40, loss: 0.045718129724264145
step: 50, loss: 0.003291957313194871
step: 60, loss: 0.007730313576757908
step: 70, loss: 0.004666922148317099
step: 80, loss: 0.00181186490226537
step: 90, loss: 0.0068716066889464855
step: 100, loss: 0.0019346955232322216
step: 110, loss: 0.0006537640583701432
step: 120, loss: 0.009745887480676174
step: 130, loss: 0.0010849963873624802
step: 140, loss: 0.12898509204387665
step: 150, loss: 0.0010601230897009373
step: 160, loss: 0.011647433042526245
step: 170, loss: 0.02677142061293125
/usr/local/lib/python3.7/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=0.681948424068768, f1=0.6927083333333334,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003659407142549753
step: 10, loss: 0.0006640874780714512
step: 20, loss: 0.0011504541616886854
step: 30, loss: 0.00500338152050972
step: 40, loss: 0.0008427802822552621
step: 50, loss: 0.0011013956973329186
step: 60, loss: 0.003721010871231556
step: 70, loss: 0.0023651740048080683
step: 80, loss: 0.003876244183629751
step: 90, loss: 0.0006001276196911931
step: 100, loss: 0.020526209846138954
step: 110, loss: 0.028972361236810684
step: 120, loss: 0.005013955757021904
step: 130, loss: 0.0012471235822886229
step: 140, loss: 0.0010426483349874616
step: 150, loss: 0.0024357738438993692
step: 160, loss: 0.0013106223195791245
step: 170, loss: 0.0007770496886223555
/usr/local/lib/python3.7/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=0.6842105263157895, f1=0.6960784313725491,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046392439980991185
step: 10, loss: 0.002737775444984436
step: 20, loss: 0.09190026670694351
step: 30, loss: 0.0009702685056254268
step: 40, loss: 0.0019111745059490204
step: 50, loss: 0.002653430914506316
step: 60, loss: 0.003672615624964237
step: 70, loss: 0.00041102594695985317
step: 80, loss: 0.004218451213091612
step: 90, loss: 0.0012539192102849483
step: 100, loss: 0.0019327295012772083
step: 110, loss: 0.013026931323111057
step: 120, loss: 0.010857226327061653
step: 130, loss: 0.0008170371875166893
step: 140, loss: 0.003802448743954301
step: 150, loss: 0.0009539689053781331
step: 160, loss: 0.001485512126237154
step: 170, loss: 0.0015258811181411147
/usr/local/lib/python3.7/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=0.6891191709844561, f1=0.7021791767554479,
best_f1=0.6812652068126521
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:09, 204.66it/s]
load_f1 = 0.5853658536585367
real_f1 = 0.5508474576271186
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 176.63it/s]

```

4.3 DITTO DIRTY

4.3.1 DBLP-ACM - Running the matcher

```

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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



```

step: 0, loss: 0.6112583875656128
step: 10, loss: 0.6425053477287292
step: 20, loss: 0.5024682283401489
step: 30, loss: 0.49299857020378113
step: 40, loss: 0.33945152163505554
step: 50, loss: 0.09824777394533157
step: 60, loss: 0.08160559833049774
step: 70, loss: 0.1350330114364624
step: 80, loss: 0.0736512690782547
step: 90, loss: 0.10158059000968933
step: 100, loss: 0.003985240124166012
step: 110, loss: 0.17033912241458893
step: 120, loss: 0.04715033248066902
step: 130, loss: 0.008792468346655369
step: 140, loss: 0.032375186681747437
step: 150, loss: 0.08091802895069122
step: 160, loss: 0.03708646446466446
step: 170, loss: 0.34693121910095215
step: 180, loss: 0.3127778172492981
step: 190, loss: 0.07206883281469345
step: 200, loss: 0.068061463534832
step: 210, loss: 0.006968977861106396
step: 220, loss: 0.015121906995773315
step: 230, loss: 0.09445174038410187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9411764705882353, f1=0.9388209121245829,
best_f1=0.9388209121245829
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054883998818695545
step: 10, loss: 0.004746670834720135
step: 20, loss: 0.2383190393447876
step: 30, loss: 0.1518149971961975
step: 40, loss: 0.1456727683544159
step: 50, loss: 0.010085509158670902
step: 60, loss: 0.015918508172035217
step: 70, loss: 0.11594723910093307
step: 80, loss: 0.06442676484584808
step: 90, loss: 0.06138131022453308
step: 100, loss: 0.12494660913944244
step: 110, loss: 0.21599242091178894

```

```

step: 120, loss: 0.1504804491996765
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.08933646976947784
step: 140, loss: 0.015073010697960854
step: 150, loss: 0.030097048729658127
step: 160, loss: 0.03942212462425232
step: 170, loss: 0.010473741218447685
step: 180, loss: 0.05961298570036888
step: 190, loss: 0.039277657866477966
step: 200, loss: 0.008935874328017235
step: 210, loss: 0.005853372626006603
step: 220, loss: 0.05233799293637276
step: 230, loss: 0.25609201192855835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9589345172031077, f1=0.9552572706935123,
best_f1=0.9552572706935123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027092942968010902
step: 10, loss: 0.05225331708788872
step: 20, loss: 0.033197950571775436
step: 30, loss: 0.022337252274155617
step: 40, loss: 0.07995156943798065
step: 50, loss: 0.06210964173078537
step: 60, loss: 0.034119658172130585
step: 70, loss: 0.02064487896859646
step: 80, loss: 0.01880904659628868
step: 90, loss: 0.03140558302402496
step: 100, loss: 0.0668257549405098
step: 110, loss: 0.0038710779044777155
step: 120, loss: 0.0069852182641625404
step: 130, loss: 0.0007567614084109664
step: 140, loss: 0.008255494758486748
step: 150, loss: 0.06366995722055435
step: 160, loss: 0.06824727356433868
step: 170, loss: 0.021875331178307533
step: 180, loss: 0.02778995782136917
step: 190, loss: 0.007403744384646416
step: 200, loss: 0.012689029797911644
step: 210, loss: 0.023953642696142197
step: 220, loss: 0.0028621105011552572

```

```

step: 230, loss: 0.0014978778781369328
/usr/local/lib/python3.7/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=0.9581497797356827, f1=0.9565217391304347,
best_f1=0.9552572706935123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002128481399267912
step: 10, loss: 0.002102195518091321
step: 20, loss: 0.00857420265674591
step: 30, loss: 0.014491681009531021
step: 40, loss: 0.01312005240470171
step: 50, loss: 0.004489723592996597
step: 60, loss: 0.003617602400481701
step: 70, loss: 0.0025545204989612103
step: 80, loss: 0.002769271144643426
step: 90, loss: 0.020934859290719032
step: 100, loss: 0.004403913859277964
step: 110, loss: 0.012806234881281853
step: 120, loss: 0.02158365771174431
step: 130, loss: 0.1281825304031372
step: 140, loss: 0.002378285164013505
step: 150, loss: 0.2585284411907196
step: 160, loss: 0.060101110488176346
step: 170, loss: 0.024404427036643028
step: 180, loss: 0.0009801206178963184
step: 190, loss: 0.0022900686599314213
step: 200, loss: 0.004465287551283836
step: 210, loss: 0.0016358408611267805
step: 220, loss: 0.001095120212994516
step: 230, loss: 0.013611667789518833
/usr/local/lib/python3.7/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=0.9565217391304347, f1=0.9511111111111112,
best_f1=0.9552572706935123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07453760504722595
step: 10, loss: 0.001501749036833644
step: 20, loss: 0.006373519543558359
step: 30, loss: 0.0007305326289497316
step: 40, loss: 0.0014573123771697283
step: 50, loss: 0.02483501099050045
step: 60, loss: 0.051665112376213074

```

```

step: 70, loss: 0.0017697729635983706
step: 80, loss: 0.0005908524617552757
step: 90, loss: 0.007639922667294741
step: 100, loss: 0.00026029208675026894
step: 110, loss: 0.0002736237074714154
step: 120, loss: 0.0003951505059376359
step: 130, loss: 0.03634939342737198
step: 140, loss: 0.005677782464772463
step: 150, loss: 0.005102585535496473
step: 160, loss: 0.0003508070658426732
step: 170, loss: 0.00901932455599308
step: 180, loss: 0.00033581574098207057
step: 190, loss: 0.11229947209358215
step: 200, loss: 0.037799522280693054
step: 210, loss: 0.011302044615149498
step: 220, loss: 0.0042775836773216724
step: 230, loss: 0.001181307015940547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9615806805708014, f1=0.9592959295929592,
best_f1=0.9592959295929592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010161417536437511
step: 10, loss: 0.0013098687632009387
step: 20, loss: 0.01041490864008665
step: 30, loss: 0.0033209885004907846
step: 40, loss: 0.00027649669209495187
step: 50, loss: 0.000373556453268975
step: 60, loss: 0.00016307497571688145
step: 70, loss: 0.0012659062631428242
step: 80, loss: 0.0005924021825194359
step: 90, loss: 0.07483331114053726
step: 100, loss: 0.13347235321998596
step: 110, loss: 0.0017763329669833183
step: 120, loss: 0.057900115847587585
step: 130, loss: 0.015267172828316689
step: 140, loss: 0.0011296411976218224
step: 150, loss: 0.02741832844913006
step: 160, loss: 0.004367985296994448
step: 170, loss: 0.007115104701370001
step: 180, loss: 0.007658190093934536

```

```

step: 190, loss: 0.057818781584501266
step: 200, loss: 0.001843275036662817
step: 210, loss: 0.0012697218917310238
step: 220, loss: 0.01334310881793499
step: 230, loss: 0.08035188168287277
/usr/local/lib/python3.7/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=0.9542920847268673, f1=0.9502262443438914,
best_f1=0.9592959295929592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012967672199010849
step: 10, loss: 0.0013730202335864305
step: 20, loss: 0.0038845252711325884
step: 30, loss: 0.02867252193391323
step: 40, loss: 0.008632884360849857
step: 50, loss: 0.00036581663880497217
step: 60, loss: 0.0008908368181437254
step: 70, loss: 0.011061522178351879
step: 80, loss: 0.00320099457167089
step: 90, loss: 0.000136314207338728
step: 100, loss: 0.00038212802610360086
step: 110, loss: 0.0017681081080809236
step: 120, loss: 0.002453394467011094
step: 130, loss: 0.004980150144547224
step: 140, loss: 0.00029323133639991283
step: 150, loss: 0.003074508858844638
step: 160, loss: 0.08970427513122559
step: 170, loss: 0.011120768263936043
step: 180, loss: 0.003750344505533576
step: 190, loss: 0.000582112988922745
step: 200, loss: 0.12755347788333893
step: 210, loss: 7.813864067429677e-05
step: 220, loss: 0.07442910224199295
step: 230, loss: 0.08824358135461807
/usr/local/lib/python3.7/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=0.9588431590656283, f1=0.953229398663697,
best_f1=0.9592959295929592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002375858457526192
step: 10, loss: 0.07718590646982193
step: 20, loss: 0.001225322368554771

```

```

step: 30, loss: 0.00473802350461483
step: 40, loss: 0.05004673823714256
step: 50, loss: 0.0007141176029108465
step: 60, loss: 0.0004542275273706764
step: 70, loss: 0.0017568678595125675
step: 80, loss: 0.0003470703086350113
step: 90, loss: 0.00020648882491514087
step: 100, loss: 0.04269363731145859
step: 110, loss: 0.0016841592732816935
step: 120, loss: 0.025312121957540512
step: 130, loss: 0.005010436289012432
step: 140, loss: 0.000435110239777714
step: 150, loss: 0.00030767059070058167
step: 160, loss: 0.001226886291988194
step: 170, loss: 0.0007747897179797292
step: 180, loss: 0.0725911557674408
step: 190, loss: 0.004548244643956423
step: 200, loss: 0.002362368628382683
step: 210, loss: 0.0022854015696793795
step: 220, loss: 0.000546671508345753
step: 230, loss: 0.09005814045667648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9683257918552037, f1=0.9477272727272728,
best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009008195484057069
step: 10, loss: 0.006636475678533316
step: 20, loss: 0.001219981350004673
step: 30, loss: 9.510829113423824e-05
step: 40, loss: 0.01686665415763855
step: 50, loss: 0.00015994218119885772
step: 60, loss: 0.0002871298638638109
step: 70, loss: 0.0001357321598334238
step: 80, loss: 8.845824049785733e-05
step: 90, loss: 0.00028026284417137504
step: 100, loss: 0.0005614759866148233
step: 110, loss: 0.0006097727455198765
step: 120, loss: 0.0019330017967149615
step: 130, loss: 0.00691326754167676
step: 140, loss: 0.0003852092777378857

```

```

step: 150, loss: 0.14178387820720673
step: 160, loss: 0.00021997479780111462
step: 170, loss: 0.0003496247227303684
step: 180, loss: 0.005569511093199253
step: 190, loss: 0.0020658313296735287
step: 200, loss: 0.000668024702463299
step: 210, loss: 0.0013546643313020468
step: 220, loss: 0.0001413434511050582
step: 230, loss: 0.0004365343484096229
/usr/local/lib/python3.7/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=0.9627959413754228, f1=0.9532497149372862,
best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023467213031835854
step: 10, loss: 0.000356759614078328
step: 20, loss: 0.004093323368579149
step: 30, loss: 8.936470840126276e-05
step: 40, loss: 0.0001241543359356001
step: 50, loss: 0.0009013499948196113
step: 60, loss: 0.0005235809949226677
step: 70, loss: 0.0002297440660186112
step: 80, loss: 7.859031757107005e-05
step: 90, loss: 0.003040907671675086
step: 100, loss: 7.571111927973107e-05
step: 110, loss: 9.122284973273054e-05
step: 120, loss: 0.0002974549133796245
step: 130, loss: 0.02381480112671852
step: 140, loss: 0.057502150535583496
step: 150, loss: 0.01424513477832079
step: 160, loss: 0.0011886009015142918
step: 170, loss: 5.298212272464298e-05
step: 180, loss: 0.00036640255711972713
step: 190, loss: 0.011015078984200954
step: 200, loss: 5.441206303657964e-05
step: 210, loss: 9.095874702325091e-05
step: 220, loss: 0.00014643445319961756
step: 230, loss: 0.05812162905931473
/usr/local/lib/python3.7/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=0.9620535714285715, f1=0.953125, best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0002557651314418763
step: 10, loss: 0.00016782335296738893
step: 20, loss: 0.0012651474680751562
step: 30, loss: 0.006614319514483213
step: 40, loss: 0.00010366813512519002
step: 50, loss: 0.00048693138523958623
step: 60, loss: 0.00011873492621816695
step: 70, loss: 0.0002597629209049046
step: 80, loss: 5.1149469072697684e-05
step: 90, loss: 6.654221215285361e-05
step: 100, loss: 7.345630001509562e-05
step: 110, loss: 0.0006068412913009524
step: 120, loss: 0.000114080285129603
step: 130, loss: 0.00019690915360115469
step: 140, loss: 0.0002706922823563218
step: 150, loss: 0.012597236782312393
step: 160, loss: 4.8743222578195855e-05
step: 170, loss: 0.0003383269358891994
step: 180, loss: 7.10313834133558e-05
step: 190, loss: 6.384873267961666e-05
step: 200, loss: 0.00019371161761227995
step: 210, loss: 3.974002174800262e-05
step: 220, loss: 5.5436492402805015e-05
step: 230, loss: 4.4235999666852877e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.961883408071749, f1=0.9534619750283768,
best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.549835911253467e-05
step: 10, loss: 0.0003852683585137129
step: 20, loss: 0.00016720571147743613
step: 30, loss: 7.767155329929665e-05
step: 40, loss: 0.07959844172000885
step: 50, loss: 0.004191108047962189
step: 60, loss: 0.0005493620410561562
step: 70, loss: 0.00011988209735136479
step: 80, loss: 9.325014252681285e-05
step: 90, loss: 6.994366412982345e-05
step: 100, loss: 4.626409281627275e-05
step: 110, loss: 0.00023880283697508276
step: 120, loss: 4.735718903248198e-05
step: 130, loss: 0.00021445401944220066
step: 140, loss: 0.00049094099085778
step: 150, loss: 0.00012147127563366666

```



```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 9.606201638234779e-05
step: 170, loss: 0.0007084383978508413
step: 180, loss: 0.00501065980643034
step: 190, loss: 0.0015563147608190775
step: 200, loss: 3.3720152714522555e-05
step: 210, loss: 5.23415110365022e-05
step: 220, loss: 0.00010230124462395906
step: 230, loss: 0.0075393542647361755
/usr/local/lib/python3.7/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=0.9584736251402918, f1=0.9466515323496028,
best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006144444341771305
step: 10, loss: 0.00010855899745365605
step: 20, loss: 0.00015476878616027534
step: 30, loss: 5.443485497380607e-05
step: 40, loss: 0.0014388726558536291
step: 50, loss: 6.871324876556173e-05
step: 60, loss: 0.0008439699886366725
step: 70, loss: 5.0108319555874914e-05
step: 80, loss: 7.22624099580571e-05
step: 90, loss: 0.0008837233181111515
step: 100, loss: 0.000133522305986844
step: 110, loss: 0.042940422892570496
step: 120, loss: 0.0004628312890417874
step: 130, loss: 0.00010289045894751325
step: 140, loss: 0.00015412690117955208
step: 150, loss: 4.705579704022966e-05
step: 160, loss: 0.00022932360298000276
step: 170, loss: 0.00020540445984806865
step: 180, loss: 0.0007001168560236692
step: 190, loss: 4.7685596655355766e-05
step: 200, loss: 0.00037845922634005547
step: 210, loss: 3.413773447391577e-05
step: 220, loss: 4.884679947281256e-05
step: 230, loss: 4.855894439970143e-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=0.9609810479375697, f1=0.9518477043673013,
best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.929003262077458e-05
step: 10, loss: 0.00014473979535978287
step: 20, loss: 0.000862862216308713
step: 30, loss: 5.783838059869595e-05
step: 40, loss: 0.006566113792359829
step: 50, loss: 0.003829807974398136
step: 60, loss: 9.552703704684973e-05
step: 70, loss: 8.009788143681362e-05
step: 80, loss: 5.3277686674846336e-05
step: 90, loss: 0.0003189978888258338
step: 100, loss: 0.0003449555370025337
step: 110, loss: 0.0019372425740584731
step: 120, loss: 3.394006489543244e-05
step: 130, loss: 0.00012562509800773114
step: 140, loss: 3.8495782064273953e-05
step: 150, loss: 3.2870855648070574e-05
step: 160, loss: 9.924601909006014e-05
step: 170, loss: 6.126508378656581e-05
step: 180, loss: 0.0010994464391842484
step: 190, loss: 3.27818124787882e-05
step: 200, loss: 7.06591090420261e-05
step: 210, loss: 6.177992327138782e-05
step: 220, loss: 0.00014880088565405458
step: 230, loss: 0.0012485483894124627
/usr/local/lib/python3.7/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=0.9615384615384616, f1=0.9519450800915331,
best_f1=0.9477272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.148540054098703e-05
step: 10, loss: 4.4409651309251785e-05
step: 20, loss: 4.2484414734644815e-05
step: 30, loss: 0.0002534427912905812
step: 40, loss: 5.680646063410677e-05
step: 50, loss: 0.0464373417198658
step: 60, loss: 4.905600144411437e-05
step: 70, loss: 0.001490571303293109
step: 80, loss: 0.00023419628269039094
step: 90, loss: 0.0003835927345789969
step: 100, loss: 4.336413985583931e-05
step: 110, loss: 3.425314207561314e-05
step: 120, loss: 0.0003442020097281784
step: 130, loss: 5.413438702817075e-05
step: 140, loss: 0.003133351681753993

```

```

step: 150, loss: 0.004276635590940714
step: 160, loss: 7.984452531673014e-05
step: 170, loss: 3.0349119697348215e-05
step: 180, loss: 9.752465120982379e-05
step: 190, loss: 0.0012140849139541388
step: 200, loss: 0.0007098286296240985
step: 210, loss: 0.0002674812276381999
step: 220, loss: 0.0020343014039099216
step: 230, loss: 5.5874646932352334e-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=0.9594594594594594, f1=0.9533560864618885,
best_f1=0.9477272727272728
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.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:16, 152.92it/s]
load_f1 = 0.9641255605381166
real_f1 = 0.9619686800894856
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:25, 175.26it/s]
```

4.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

```
Downloading: 100% 433/433 [00:00<00:00, 450kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 1.23MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 69.7MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.

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

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

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

step: 0, loss: 0.6308624744415283

step: 10, loss: 0.6032471656799316

step: 20, loss: 0.5994514226913452

step: 30, loss: 0.3534665107727051

step: 40, loss: 0.3783387839794159

step: 50, loss: 0.22612157464027405

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: 60, loss: 0.10589089244604111

step: 70, loss: 0.2236054241657257

step: 80, loss: 0.10979318618774414

step: 90, loss: 0.40887370705604553

step: 100, loss: 0.19607095420360565

step: 110, loss: 0.129455104470253

step: 120, loss: 0.11008865386247635

```
step: 130, loss: 0.179279625415802
step: 140, loss: 0.11163455992937088
step: 150, loss: 0.09006695449352264
step: 160, loss: 0.1105445921421051
step: 170, loss: 0.30319690704345703
step: 180, loss: 0.08209290355443954
step: 190, loss: 0.019269302487373352
step: 200, loss: 0.15609294176101685
step: 210, loss: 0.060745444148778915
step: 220, loss: 0.2706945538520813
step: 230, loss: 0.1904507726430893
step: 240, loss: 0.08724727481603622
step: 250, loss: 0.10056265443563461
step: 260, loss: 0.06593586504459381
step: 270, loss: 0.023828720673918724
step: 280, loss: 0.19308306276798248
step: 290, loss: 0.2809520363807678
step: 300, loss: 0.06415525078773499
step: 310, loss: 0.23609450459480286
step: 320, loss: 0.05800242722034454
step: 330, loss: 0.09501051157712936
step: 340, loss: 0.21608494222164154
step: 350, loss: 0.046079427003860474
step: 360, loss: 0.0917193815112114
step: 370, loss: 0.10223643481731415
step: 380, loss: 0.011865193955600262
step: 390, loss: 0.07146729528903961
step: 400, loss: 0.3023937940597534
step: 410, loss: 0.06558353453874588
step: 420, loss: 0.09069307893514633
step: 430, loss: 0.1747521013021469
step: 440, loss: 0.09228198230266571
step: 450, loss: 0.06431342661380768
step: 460, loss: 0.00969283189624548
step: 470, loss: 0.1327700912952423
step: 480, loss: 0.07971636950969696
step: 490, loss: 0.0853620246052742
step: 500, loss: 0.08713164180517197
step: 510, loss: 0.11188672482967377
step: 520, loss: 0.13075077533721924
step: 530, loss: 0.008883076719939709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9072261072261073, f1=0.8876351669017396,  
best_f1=0.8876351669017396  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.18230444192886353  
step: 10, loss: 0.14306297898292542  
step: 20, loss: 0.20137865841388702  
step: 30, loss: 0.08887867629528046  
step: 40, loss: 0.06548857688903809  
step: 50, loss: 0.11013713479042053  
step: 60, loss: 0.07806499302387238  
step: 70, loss: 0.048743657767772675  
step: 80, loss: 0.06817272305488586  
step: 90, loss: 0.07117559760808945  
step: 100, loss: 0.09879474341869354  
step: 110, loss: 0.10872165858745575  
step: 120, loss: 0.08267752081155777  
step: 130, loss: 0.11053754389286041  
step: 140, loss: 0.0682210773229599  
step: 150, loss: 0.13777834177017212  
step: 160, loss: 0.023302676156163216  
step: 170, loss: 0.13033680617809296  
step: 180, loss: 0.0454958900809288  
step: 190, loss: 0.06711029261350632  
step: 200, loss: 0.029953598976135254  
step: 210, loss: 0.086685411632061  
step: 220, loss: 0.08574916422367096  
step: 230, loss: 0.05620792880654335  
step: 240, loss: 0.061646413058042526  
step: 250, loss: 0.10141347348690033  
step: 260, loss: 0.051630258560180664  
step: 270, loss: 0.1753477156162262  
step: 280, loss: 0.16053202748298645  
step: 290, loss: 0.02872966229915619  
step: 300, loss: 0.2454231083393097  
step: 310, loss: 0.04407922923564911  
step: 320, loss: 0.1952347457408905  
step: 330, loss: 0.09094320982694626  
step: 340, loss: 0.07831069082021713  
step: 350, loss: 0.015131842344999313  
step: 360, loss: 0.148757204413414  
step: 370, loss: 0.1635657250881195  
step: 380, loss: 0.11317852139472961  
step: 390, loss: 0.1463436633348465  
step: 400, loss: 0.01757032424211502  
step: 410, loss: 0.06074519827961922  
step: 420, loss: 0.02724740095436573
```

```

step: 430, loss: 0.15838950872421265
step: 440, loss: 0.08900714665651321
step: 450, loss: 0.1686776876449585
step: 460, loss: 0.02454710565507412
step: 470, loss: 0.02230921946465969
step: 480, loss: 0.18457531929016113
step: 490, loss: 0.016655782237648964
step: 500, loss: 0.3514931797981262
step: 510, loss: 0.03749142214655876
step: 520, loss: 0.10233378410339355
step: 530, loss: 0.10562456399202347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9101019462465245, f1=0.8976377952755905,
best_f1=0.8976377952755905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09090369194746017
step: 10, loss: 0.05051650479435921
step: 20, loss: 0.14046929776668549
step: 30, loss: 0.2116367518901825
step: 40, loss: 0.05286232382059097
step: 50, loss: 0.16880419850349426
step: 60, loss: 0.07688988745212555
step: 70, loss: 0.016828911378979683
step: 80, loss: 0.017465464770793915
step: 90, loss: 0.03802984580397606
step: 100, loss: 0.01624738983809948
step: 110, loss: 0.004249974153935909
step: 120, loss: 0.02563210017979145
step: 130, loss: 0.0025322865694761276
step: 140, loss: 0.012266874313354492
step: 150, loss: 0.1898767352104187
step: 160, loss: 0.005419933702796698
step: 170, loss: 0.132664293050766
step: 180, loss: 0.028715714812278748
step: 190, loss: 0.025567667558789253
step: 200, loss: 0.061706580221652985
step: 210, loss: 0.0473773367702961
step: 220, loss: 0.11505324393510818
step: 230, loss: 0.2111319899559021
step: 240, loss: 0.14227616786956787

```



```

step: 250, loss: 0.008944236673414707
step: 260, loss: 0.07475295662879944
step: 270, loss: 0.022482501342892647
step: 280, loss: 0.2703956365585327
step: 290, loss: 0.00650831637904048
step: 300, loss: 0.05522974207997322
step: 310, loss: 0.0450286865234375
step: 320, loss: 0.04817254841327667
step: 330, loss: 0.00792610365897417
step: 340, loss: 0.013992128893733025
step: 350, loss: 0.03602888435125351
step: 360, loss: 0.025534575805068016
step: 370, loss: 0.016985880210995674
step: 380, loss: 0.002880462910979986
step: 390, loss: 0.015180072747170925
step: 400, loss: 0.08383985608816147
step: 410, loss: 0.03227343410253525
step: 420, loss: 0.14809182286262512
step: 430, loss: 0.010795806534588337
step: 440, loss: 0.07860787212848663
step: 450, loss: 0.051675789058208466
step: 460, loss: 0.044970132410526276
step: 470, loss: 0.0541149377822876
step: 480, loss: 0.04470234364271164
step: 490, loss: 0.005834745708853006
step: 500, loss: 0.11362612992525101
step: 510, loss: 0.011050084605813026
step: 520, loss: 0.05607800558209419
step: 530, loss: 0.12204436212778091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9153992395437264, f1=0.8953655040611561,
best_f1=0.8953655040611561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0066430349834263325
step: 10, loss: 0.027809735387563705
step: 20, loss: 0.004516819026321173
step: 30, loss: 0.011271648108959198
step: 40, loss: 0.0685976967215538
step: 50, loss: 0.01387912780046463
step: 60, loss: 0.013557527214288712

```

step: 70, loss: 0.02534766122698784
step: 80, loss: 0.004326429218053818
step: 90, loss: 0.15591412782669067
step: 100, loss: 0.019829824566841125
step: 110, loss: 0.0886169895529747
step: 120, loss: 0.00245283218100667
step: 130, loss: 0.0021528112702071667
step: 140, loss: 0.004986746236681938
step: 150, loss: 0.046438705176115036
step: 160, loss: 0.0035173564683645964
step: 170, loss: 0.005279794801026583
step: 180, loss: 0.01854580268263817
step: 190, loss: 0.086207315325737
step: 200, loss: 0.02681104652583599
step: 210, loss: 0.04812709614634514
step: 220, loss: 0.007310991641134024
step: 230, loss: 0.008267893455922604
step: 240, loss: 0.0021908998023718596
step: 250, loss: 0.06312011182308197
step: 260, loss: 0.010400092229247093
step: 270, loss: 0.061282768845558167
step: 280, loss: 0.15071019530296326
step: 290, loss: 0.012388370931148529
step: 300, loss: 0.0014819566858932376
step: 310, loss: 0.006693318020552397
step: 320, loss: 0.0378216952085495
step: 330, loss: 0.0630405992269516
step: 340, loss: 0.2916943430900574
step: 350, loss: 0.04358921945095062
step: 360, loss: 0.07297121733427048
step: 370, loss: 0.003690264653414488
step: 380, loss: 0.026814214885234833
step: 390, loss: 0.03694860264658928
step: 400, loss: 0.001381085254251957
step: 410, loss: 0.02422146126627922
step: 420, loss: 0.23613262176513672
step: 430, loss: 0.0866832360625267
step: 440, loss: 0.004209062084555626
step: 450, loss: 0.017103102058172226
step: 460, loss: 0.005303812213242054
step: 470, loss: 0.039907604455947876
step: 480, loss: 0.023829763755202293
step: 490, loss: 0.019351962953805923
step: 500, loss: 0.012563557364046574
step: 510, loss: 0.09942575544118881
step: 520, loss: 0.09480052441358566
step: 530, loss: 0.042420394718647

/usr/local/lib/python3.7/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=0.9087574507106833, f1=0.9040333796940195,  
best_f1=0.8953655040611561  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.19698697328567505  
step: 10, loss: 0.04114644601941109  
step: 20, loss: 0.05112839490175247  
step: 30, loss: 0.003441245760768652  
step: 40, loss: 0.1142912432551384  
step: 50, loss: 0.001889569335617125  
step: 60, loss: 0.14990022778511047  
step: 70, loss: 0.05975887551903725  
step: 80, loss: 0.07479606568813324  
step: 90, loss: 0.0048789335414767265  
step: 100, loss: 0.007001224905252457  
step: 110, loss: 0.0009265068219974637  
step: 120, loss: 0.01891174353659153  
step: 130, loss: 0.0017293645069003105  
step: 140, loss: 0.00674563879147172  
step: 150, loss: 0.0549921877682209  
step: 160, loss: 0.0022861238103359938  
step: 170, loss: 0.007045604754239321  
step: 180, loss: 0.004460544791072607  
step: 190, loss: 0.01654401421546936  
step: 200, loss: 0.022402871400117874  
step: 210, loss: 0.037443362176418304  
step: 220, loss: 0.010975463315844536  
step: 230, loss: 0.026925817131996155  
step: 240, loss: 0.012880288995802402  
step: 250, loss: 0.014495009556412697  
step: 260, loss: 0.005816080141812563  
step: 270, loss: 0.012859560549259186  
step: 280, loss: 0.0021992544643580914  
step: 290, loss: 0.236911341547966  
step: 300, loss: 0.034711651504039764  
step: 310, loss: 0.009963475167751312  
step: 320, loss: 0.1562647968530655  
step: 330, loss: 0.046950288116931915  
step: 340, loss: 0.0026724543422460556  
step: 350, loss: 0.0008621826418675482  
step: 360, loss: 0.005734287202358246  
step: 370, loss: 0.014865242876112461  
step: 380, loss: 0.0011756597086787224  
step: 390, loss: 0.000926606822758913  
step: 400, loss: 0.03421005606651306
```

```

step: 410, loss: 0.014734147116541862
step: 420, loss: 0.0012469551293179393
step: 430, loss: 0.10551777482032776
step: 440, loss: 0.17245620489120483
step: 450, loss: 0.00886040274053812
step: 460, loss: 0.0050578126683831215
step: 470, loss: 0.005866053979843855
step: 480, loss: 0.03171348571777344
step: 490, loss: 0.001454837154597044
step: 500, loss: 0.07025256007909775
step: 510, loss: 0.0030387442093342543
step: 520, loss: 0.03448103740811348
step: 530, loss: 0.009390738792717457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9168207024029574, f1=0.8997214484679666,
best_f1=0.8997214484679666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015101785538718104
step: 10, loss: 0.002478862414136529
step: 20, loss: 0.11003152281045914
step: 30, loss: 0.006985130254179239
step: 40, loss: 0.0017004329711198807
step: 50, loss: 0.0033593117259442806
step: 60, loss: 0.0001971826422959566
step: 70, loss: 0.0012164624640718102
step: 80, loss: 0.0514473021030426
step: 90, loss: 0.005910306703299284
step: 100, loss: 0.023439977318048477
step: 110, loss: 0.011378182098269463
step: 120, loss: 0.024989865720272064
step: 130, loss: 0.001539392862468958
step: 140, loss: 0.021524429321289062
step: 150, loss: 0.006504649296402931
step: 160, loss: 0.002761795185506344
step: 170, loss: 0.007358884904533625
step: 180, loss: 0.002614314202219248
step: 190, loss: 0.006294431164860725
step: 200, loss: 0.030139822512865067
step: 210, loss: 0.0018709105206653476
step: 220, loss: 0.0027594794519245625

```

```

step: 230, loss: 0.004641438368707895
step: 240, loss: 0.05480504408478737
step: 250, loss: 0.004973806440830231
step: 260, loss: 0.012785669416189194
step: 270, loss: 0.014854823239147663
step: 280, loss: 0.001810629852116108
step: 290, loss: 0.0004599808598868549
step: 300, loss: 0.00039926698082126677
step: 310, loss: 0.0013919607736170292
step: 320, loss: 0.0008702654740773141
step: 330, loss: 0.004096830729395151
step: 340, loss: 0.07715679705142975
step: 350, loss: 0.024414919316768646
step: 360, loss: 0.04565298929810524
step: 370, loss: 0.0013741697184741497
step: 380, loss: 0.0007200422114692628
step: 390, loss: 0.11618383973836899
step: 400, loss: 0.11270009726285934
step: 410, loss: 0.007898246869444847
step: 420, loss: 0.056369464844465256
step: 430, loss: 0.001282812561839819
step: 440, loss: 0.005339452065527439
step: 450, loss: 0.0013114108005538583
step: 460, loss: 0.008757714182138443
step: 470, loss: 0.016695402562618256
step: 480, loss: 0.022279368713498116
step: 490, loss: 0.003732894780114293
step: 500, loss: 0.0008756423485465348
step: 510, loss: 0.007275558542460203
step: 520, loss: 0.004005839116871357
step: 530, loss: 0.019645147025585175
/usr/local/lib/python3.7/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=0.9031352363125877, f1=0.8772426817752597,
best_f1=0.8997214484679666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031515348702669144
step: 10, loss: 0.00924545805901289
step: 20, loss: 0.17961455881595612
step: 30, loss: 0.00861323717981577
step: 40, loss: 0.03483874723315239
step: 50, loss: 0.014020106755197048
step: 60, loss: 0.011027703993022442
step: 70, loss: 0.0010788252111524343
step: 80, loss: 0.016148356720805168

```

```
step: 90, loss: 0.0011790626449510455
step: 100, loss: 0.00013245080481283367
step: 110, loss: 0.00038339311140589416
step: 120, loss: 0.0016973029123619199
step: 130, loss: 0.016889410093426704
step: 140, loss: 0.00032386419479735196
step: 150, loss: 0.00275716301985085
step: 160, loss: 0.09211436659097672
step: 170, loss: 0.002797223860397935
step: 180, loss: 0.0010187311563640833
step: 190, loss: 0.13661330938339233
step: 200, loss: 0.010134467855095863
step: 210, loss: 0.00695136608555913
step: 220, loss: 0.0009410791099071503
step: 230, loss: 0.029287273064255714
step: 240, loss: 0.0025969459675252438
step: 250, loss: 0.0004650845075957477
step: 260, loss: 0.004686142783612013
step: 270, loss: 0.030994830653071404
step: 280, loss: 0.10990392416715622
step: 290, loss: 0.004669333342462778
step: 300, loss: 0.025307755917310715
step: 310, loss: 0.0004367638612166047
step: 320, loss: 0.0008648757939226925
step: 330, loss: 0.07851580530405045
step: 340, loss: 0.0026654438115656376
step: 350, loss: 0.0005350107094272971
step: 360, loss: 0.002436259062960744
step: 370, loss: 0.0021530338563024998
step: 380, loss: 0.0016151946038007736
step: 390, loss: 0.0005999439745210111
step: 400, loss: 0.0037773705553263426
step: 410, loss: 0.0017030773451551795
step: 420, loss: 0.010874301195144653
step: 430, loss: 0.0028280564583837986
step: 440, loss: 0.0012089544907212257
step: 450, loss: 0.002166761551052332
step: 460, loss: 0.0022217051591724157
step: 470, loss: 0.03783382102847099
step: 480, loss: 0.2232976108789444
step: 490, loss: 0.006647947244346142
step: 500, loss: 0.005814340431243181
step: 510, loss: 0.007022024132311344
step: 520, loss: 0.12662021815776825
step: 530, loss: 0.037307221442461014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.917397323488694, f1=0.9006010171058715,
best_f1=0.9006010171058715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013784898445010185
step: 10, loss: 0.01779925636947155
step: 20, loss: 0.00038212264189496636
step: 30, loss: 0.001509202760644257
step: 40, loss: 0.000593398988712579
step: 50, loss: 0.02469637617468834
step: 60, loss: 0.0007074191817082465
step: 70, loss: 0.0004301226290408522
step: 80, loss: 0.00383172370493412
step: 90, loss: 0.0030528386123478413
step: 100, loss: 0.0627811998128891
step: 110, loss: 0.0018292561871930957
step: 120, loss: 0.04289949685335159
step: 130, loss: 0.001766806934028864
step: 140, loss: 0.01070067472755909
step: 150, loss: 0.0004042889049742371
step: 160, loss: 0.0051720160990953445
step: 170, loss: 0.020923800766468048
step: 180, loss: 0.006710372865200043
step: 190, loss: 0.002775453496724367
step: 200, loss: 0.001840410870499909
step: 210, loss: 0.008254650980234146
step: 220, loss: 0.005098004825413227
step: 230, loss: 0.0497889444231987
step: 240, loss: 0.0007880938937887549
step: 250, loss: 0.008777955546975136
step: 260, loss: 0.05641230195760727
step: 270, loss: 0.04136605188250542
step: 280, loss: 0.07160801440477371
step: 290, loss: 0.016888462007045746
step: 300, loss: 0.014897353947162628
step: 310, loss: 0.002064456231892109
step: 320, loss: 0.22029389441013336
step: 330, loss: 0.0033292328007519245
step: 340, loss: 0.003123251721262932
step: 350, loss: 0.0018908934434875846
step: 360, loss: 0.007102487608790398
step: 370, loss: 0.004523665644228458
step: 380, loss: 0.004714195150882006
```

```

step: 390, loss: 0.0136777488514781
step: 400, loss: 0.00027766742277890444
step: 410, loss: 0.0002673873968888074
step: 420, loss: 0.001377498498186469
step: 430, loss: 0.02418537251651287
step: 440, loss: 0.0034636922646313906
step: 450, loss: 0.0002074568037642166
step: 460, loss: 0.002977213589474559
step: 470, loss: 0.0004710783832706511
step: 480, loss: 0.0006142063066363335
step: 490, loss: 0.09853112697601318
step: 500, loss: 0.0006809444166719913
step: 510, loss: 0.006708059925585985
step: 520, loss: 0.004592966288328171
step: 530, loss: 0.003128133248537779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9235048678720444, f1=0.8941724941724941,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029088398441672325
step: 10, loss: 0.0009015353280119598
step: 20, loss: 0.09525834769010544
step: 30, loss: 0.0015663934173062444
step: 40, loss: 0.0015379594406113029
step: 50, loss: 0.0023508816957473755
step: 60, loss: 0.0003483761101961136
step: 70, loss: 0.00021664160885848105
step: 80, loss: 0.002188457176089287
step: 90, loss: 0.00036889046896249056
step: 100, loss: 0.0004609638126567006
step: 110, loss: 0.000688511470798403
step: 120, loss: 0.01812032423913479
step: 130, loss: 0.10832741111516953
step: 140, loss: 0.04985828697681427
step: 150, loss: 0.022276924923062325
step: 160, loss: 0.002752905711531639
step: 170, loss: 0.0023855865001678467
step: 180, loss: 0.0009531673276796937
step: 190, loss: 0.0063118417747318745
step: 200, loss: 0.0012989339884370565

```



```

step: 210, loss: 0.0022688594181090593
step: 220, loss: 0.003243141109123826
step: 230, loss: 0.004197880160063505
step: 240, loss: 0.0005717940512113273
step: 250, loss: 0.0005813332390971482
step: 260, loss: 0.00856807641685009
step: 270, loss: 0.000510913785547018
step: 280, loss: 0.0020644140895456076
step: 290, loss: 0.07220793515443802
step: 300, loss: 0.00012951558164786547
step: 310, loss: 0.000771035032812506
step: 320, loss: 0.000324392196489498
step: 330, loss: 0.00612950325012207
step: 340, loss: 0.0019810288213193417
step: 350, loss: 0.008640659973025322
step: 360, loss: 0.00042651043622754514
step: 370, loss: 0.002063347725197673
step: 380, loss: 0.0005362993688322604
step: 390, loss: 0.0035701736342161894
step: 400, loss: 0.03143167123198509
step: 410, loss: 0.001307621831074357
step: 420, loss: 0.0011962155112996697
step: 430, loss: 0.0006536361761391163
step: 440, loss: 0.0019401764729991555
step: 450, loss: 0.00045146053889766335
step: 460, loss: 0.00021910197392571718
step: 470, loss: 0.0005798537749797106
step: 480, loss: 0.04346496984362602
step: 490, loss: 0.0009246842819266021
step: 500, loss: 0.0001871815911727026
step: 510, loss: 0.002310340292751789
step: 520, loss: 0.0006381258717738092
step: 530, loss: 0.00015077581338118762
/usr/local/lib/python3.7/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=0.9131423374261028, f1=0.8932481751824817,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010257872054353356
step: 10, loss: 0.00010399777966085821
step: 20, loss: 4.512314262683503e-05
step: 30, loss: 5.309411426424049e-05
step: 40, loss: 0.00010022974311141297
step: 50, loss: 0.000108095642644912
step: 60, loss: 0.0034215631894767284

```

step: 70, loss: 0.00018166577501688153
step: 80, loss: 0.0020355554297566414
step: 90, loss: 0.0010147785069420934
step: 100, loss: 0.00015119221643544734
step: 110, loss: 0.0001845216320361942
step: 120, loss: 0.0001661337009863928
step: 130, loss: 0.12027030438184738
step: 140, loss: 0.001913155079819262
step: 150, loss: 0.018109699711203575
step: 160, loss: 0.0019218545639887452
step: 170, loss: 0.00047687769983895123
step: 180, loss: 0.04735395312309265
step: 190, loss: 0.0004010460979770869
step: 200, loss: 0.0003213545132894069
step: 210, loss: 0.0007547424756921828
step: 220, loss: 0.006733099929988384
step: 230, loss: 0.002223188290372491
step: 240, loss: 0.003309824038296938
step: 250, loss: 0.00020722050976473838
step: 260, loss: 0.0018644254887476563
step: 270, loss: 0.0002058974641840905
step: 280, loss: 0.0002100480196531862
step: 290, loss: 0.005449139513075352
step: 300, loss: 0.00019663793500512838
step: 310, loss: 0.00024777062935754657
step: 320, loss: 0.006685202941298485
step: 330, loss: 0.0015415396774187684
step: 340, loss: 0.009023305028676987
step: 350, loss: 0.0023337334860116243
step: 360, loss: 0.009630984626710415
step: 370, loss: 0.0023325509391725063
step: 380, loss: 0.001957076136022806
step: 390, loss: 0.008733408525586128
step: 400, loss: 0.008049550466239452
step: 410, loss: 0.0035800342448055744
step: 420, loss: 0.0009894933318719268
step: 430, loss: 0.0015803711721673608
step: 440, loss: 0.04016588628292084
step: 450, loss: 0.00036614100099541247
step: 460, loss: 0.0023618037812411785
step: 470, loss: 0.0013918188633397222
step: 480, loss: 0.0008047963492572308
step: 490, loss: 0.005022969096899033
step: 500, loss: 0.008903506211936474
step: 510, loss: 0.00030753089231438935
step: 520, loss: 0.0010081450454890728
step: 530, loss: 0.0005834268522448838
/usr/local/lib/python3.7/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=0.9124594719777674, f1=0.890125173852573,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00037755409721285105
step: 10, loss: 0.27231255173683167
step: 20, loss: 0.007426001131534576
step: 30, loss: 0.0060399253852665424
step: 40, loss: 0.0038399354089051485
step: 50, loss: 0.02993875928223133
step: 60, loss: 0.0014773056609556079
step: 70, loss: 0.0018438119441270828
step: 80, loss: 0.0011974438093602657
step: 90, loss: 0.001957011641934514
step: 100, loss: 0.0008820410002954304
step: 110, loss: 0.005000808276236057
step: 120, loss: 0.00016521233192179352
step: 130, loss: 0.00027885293820872903
step: 140, loss: 0.0035246426705271006
step: 150, loss: 0.0017888364382088184
step: 160, loss: 0.002046645851805806
step: 170, loss: 0.00031374243553727865
step: 180, loss: 0.003954057116061449
step: 190, loss: 0.008642596192657948
step: 200, loss: 0.021221281960606575
step: 210, loss: 0.0005235313437879086
step: 220, loss: 0.0002519477275200188
step: 230, loss: 7.049251871649176e-05
step: 240, loss: 0.00012108738883398473
step: 250, loss: 8.690361573826522e-05
step: 260, loss: 0.00012172155402367935
step: 270, loss: 0.19901826977729797
step: 280, loss: 0.0007142307003960013
step: 290, loss: 0.001350595150142908
step: 300, loss: 0.006945055443793535
step: 310, loss: 0.001114160637371242
step: 320, loss: 0.00032023267704062164
step: 330, loss: 0.0017730375984683633
step: 340, loss: 0.0026577136013656855
step: 350, loss: 0.014084292575716972
step: 360, loss: 0.00012153664283687249
step: 370, loss: 0.00157071347348392
step: 380, loss: 0.0001135811471613124
step: 390, loss: 4.055532917845994e-05
step: 400, loss: 8.761355275055394e-05

```

```

step: 410, loss: 0.00017235707491636276
step: 420, loss: 0.0007681387360207736
step: 430, loss: 4.582901601679623e-05
step: 440, loss: 0.00033131238888017833
step: 450, loss: 5.637381400447339e-05
step: 460, loss: 0.12904441356658936
step: 470, loss: 0.003273898968473077
step: 480, loss: 0.0032530655153095722
step: 490, loss: 0.002731436863541603
step: 500, loss: 0.0007328545907512307
step: 510, loss: 0.12394294142723083
step: 520, loss: 0.00039301745709963143
step: 530, loss: 0.00010964859393425286
/usr/local/lib/python3.7/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=0.9061488673139159, f1=0.8818897637795275,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0183677077293396
step: 10, loss: 0.00026460603112354875
step: 20, loss: 0.00011333671864122152
step: 30, loss: 0.0009893412934616208
step: 40, loss: 0.0004462438228074461
step: 50, loss: 0.02249986305832863
step: 60, loss: 0.0002675138530321419
step: 70, loss: 0.11557510495185852
step: 80, loss: 0.0005569180939346552
step: 90, loss: 0.00017982367717195302
step: 100, loss: 0.001732091885060072
step: 110, loss: 0.001060761627741158
step: 120, loss: 6.379465048667043e-05
step: 130, loss: 8.088488539215177e-05
step: 140, loss: 9.967045480152592e-05
step: 150, loss: 0.00021504929463844746
step: 160, loss: 0.0016970803262665868
step: 170, loss: 0.00045572035014629364
step: 180, loss: 0.00035324000054970384
step: 190, loss: 0.0011763775255531073
step: 200, loss: 9.850117930909619e-05
step: 210, loss: 0.0001291071966988966
step: 220, loss: 0.06942345201969147
step: 230, loss: 0.0002399972581770271
step: 240, loss: 0.0001476878533139825
step: 250, loss: 0.002188286278396845
step: 260, loss: 0.000536278763320297

```

```

step: 270, loss: 0.03210017830133438
step: 280, loss: 0.00037918216548860073
step: 290, loss: 0.00068616186035797
step: 300, loss: 0.029231050983071327
step: 310, loss: 0.00042598749860189855
step: 320, loss: 0.00028983812080696225
step: 330, loss: 0.0026251371018588543
step: 340, loss: 0.0005245704087428749
step: 350, loss: 0.0004698276170529425
step: 360, loss: 0.003944427240639925
step: 370, loss: 0.0008925388101488352
step: 380, loss: 0.0001467162510380149
step: 390, loss: 0.0004724547907244414
step: 400, loss: 0.00012032182712573558
step: 410, loss: 0.00041155581129714847
step: 420, loss: 0.002150849439203739
step: 430, loss: 0.0034478334710001945
step: 440, loss: 0.002781730378046632
step: 450, loss: 0.0006822791765443981
step: 460, loss: 9.992014383897185e-05
step: 470, loss: 0.01588597148656845
step: 480, loss: 0.001401621731929481
step: 490, loss: 0.00020952885097358376
step: 500, loss: 0.00032728808582760394
step: 510, loss: 0.05893156677484512
step: 520, loss: 5.769269046140835e-05
step: 530, loss: 0.0014549592742696404
/usr/local/lib/python3.7/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=0.909763998149005, f1=0.8893023255813953,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06833478063344955
step: 10, loss: 0.004271986428648233
step: 20, loss: 0.016468379646539688
step: 30, loss: 0.0010633267229422927
step: 40, loss: 0.00164611276704818
step: 50, loss: 0.0009568171808496118
step: 60, loss: 2.9749020541203208e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 70, loss: 0.00023852403683122247
step: 80, loss: 0.00011745314986910671
step: 90, loss: 7.625503349117935e-05
step: 100, loss: 0.0009174743900075555
step: 110, loss: 6.159744953038171e-05

```

```
step: 120, loss: 0.0006937713478691876
step: 130, loss: 0.003295470727607608
step: 140, loss: 0.0003897121059708297
step: 150, loss: 0.0006035657715983689
step: 160, loss: 0.00013302733714226633
step: 170, loss: 0.0003979481989517808
step: 180, loss: 7.439805631292984e-05
step: 190, loss: 0.002529309131205082
step: 200, loss: 0.00011909309978364035
step: 210, loss: 0.00011783946683863178
step: 220, loss: 0.0005567393964156508
step: 230, loss: 0.0001968245196621865
step: 240, loss: 9.632627188693732e-05
step: 250, loss: 5.1131068175891414e-05
step: 260, loss: 0.0004285505274310708
step: 270, loss: 0.0007450886769220233
step: 280, loss: 0.000648795859888196
step: 290, loss: 0.028241081163287163
step: 300, loss: 0.00045757999760098755
step: 310, loss: 0.002126437844708562
step: 320, loss: 0.0033274665474891663
step: 330, loss: 0.0006178246694616973
step: 340, loss: 0.00022915517911314964
step: 350, loss: 0.0003452032105997205
step: 360, loss: 4.7568213631166145e-05
step: 370, loss: 0.00047120993258431554
step: 380, loss: 0.00018473199452273548
step: 390, loss: 0.015500592067837715
step: 400, loss: 0.00074092234717682
step: 410, loss: 0.0008004317642189562
step: 420, loss: 6.513725384138525e-05
step: 430, loss: 0.004287296906113625
step: 440, loss: 0.0018348200246691704
step: 450, loss: 0.006240553688257933
step: 460, loss: 0.0007981153903529048
step: 470, loss: 0.00010040039342129603
step: 480, loss: 0.00020030456653330475
step: 490, loss: 0.00032437534537166357
step: 500, loss: 8.193973189918324e-05
step: 510, loss: 0.0006008294294588268
step: 520, loss: 0.0023930019233375788
step: 530, loss: 4.487158730626106e-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=0.9158964879852127, f1=0.8982259570494864,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 4.6482244215440005e-05
step: 10, loss: 9.438127744942904e-05
step: 20, loss: 0.0015871721552684903
step: 30, loss: 0.00010277400724589825
step: 40, loss: 0.0002213238039985299
step: 50, loss: 0.000253525358857587
step: 60, loss: 1.8510610971134156e-05
step: 70, loss: 0.0015108126681298018
step: 80, loss: 5.3952462621964514e-05
step: 90, loss: 0.00021926139015704393
step: 100, loss: 0.01707920990884304
step: 110, loss: 5.1742481446126476e-05
step: 120, loss: 6.965984357520938e-05
step: 130, loss: 0.0022292854264378548
step: 140, loss: 0.0002840344386640936
step: 150, loss: 4.144493141211569e-05
step: 160, loss: 4.1690669604577124e-05
step: 170, loss: 7.442457717843354e-05
step: 180, loss: 5.5552038247697055e-05
step: 190, loss: 0.00021228192781563848
step: 200, loss: 0.0008381931693293154
step: 210, loss: 0.004098972305655479
step: 220, loss: 0.00011191319208592176
step: 230, loss: 0.1585128754377365
step: 240, loss: 5.8643639931688085e-05
step: 250, loss: 0.00032632696093060076
step: 260, loss: 5.358775888453238e-05
step: 270, loss: 0.00034917614539153874
step: 280, loss: 0.00010665607260307297
step: 290, loss: 0.002153924899175763
step: 300, loss: 0.0016135333571583033
step: 310, loss: 0.052411846816539764
step: 320, loss: 0.000719848379958421
step: 330, loss: 0.0048605226911604404
step: 340, loss: 8.194445399567485e-05
step: 350, loss: 0.0027536824345588684
step: 360, loss: 0.0001228743785759434
step: 370, loss: 0.00040597940096631646
step: 380, loss: 0.055611129850149155
step: 390, loss: 0.0032137997914105654
step: 400, loss: 0.0011725027579814196
step: 410, loss: 0.0005324971862137318
step: 420, loss: 7.2865019319579e-05
step: 430, loss: 0.00026119116228073835
step: 440, loss: 0.001325343968346715
step: 450, loss: 0.005449975840747356
```

```

step: 460, loss: 0.00015180899936240166
step: 470, loss: 0.00022115834872238338
step: 480, loss: 0.003466630121693015
step: 490, loss: 0.000134655085275881
step: 500, loss: 0.0001470384158892557
step: 510, loss: 0.00249160616658628
step: 520, loss: 4.955637996317819e-05
step: 530, loss: 4.9769550969358534e-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=0.9137291280148423, f1=0.8987400839944004,
best_f1=0.8941724941724941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.25483483215794e-05
step: 10, loss: 0.002303576096892357
step: 20, loss: 0.00016271760978270322
step: 30, loss: 0.0009357495582662523
step: 40, loss: 0.00044786828220821917
step: 50, loss: 0.001170504023320973
step: 60, loss: 3.531221591401845e-05
step: 70, loss: 0.00027194805443286896
step: 80, loss: 0.00030651851557195187
step: 90, loss: 0.00018460489809513092
step: 100, loss: 4.348172296886332e-05
step: 110, loss: 0.0013057102914899588
step: 120, loss: 4.873567013419233e-05
step: 130, loss: 0.021505389362573624
step: 140, loss: 3.8727161154383793e-05
step: 150, loss: 0.0004440412449184805
step: 160, loss: 6.736896466463804e-05
step: 170, loss: 0.000477478199172765
step: 180, loss: 5.838947254233062e-05
step: 190, loss: 0.00027051212964579463
step: 200, loss: 0.00019622119725681841
step: 210, loss: 0.00015874276868999004
step: 220, loss: 0.0023201461881399155
step: 230, loss: 0.00046612470760010183
step: 240, loss: 0.0006141237099654973
step: 250, loss: 8.749694097787142e-05
step: 260, loss: 3.719107553479262e-05
step: 270, loss: 0.00018858499242924154
step: 280, loss: 0.00010193981870543212
step: 290, loss: 1.9572335077100433e-05
step: 300, loss: 0.00010173994087381288
step: 310, loss: 0.0032888895366340876

```



```

step: 320, loss: 0.0006662346422672272
step: 330, loss: 4.6704430133104324e-05
step: 340, loss: 8.16095998743549e-05
step: 350, loss: 4.500229624682106e-05
step: 360, loss: 0.001224661129526794
step: 370, loss: 3.883178942487575e-05
step: 380, loss: 0.00015693434397689998
step: 390, loss: 0.00019026707741431892
step: 400, loss: 0.0014338534092530608
step: 410, loss: 0.009933257475495338
step: 420, loss: 0.00018162875494454056
step: 430, loss: 0.0016426292713731527
step: 440, loss: 0.00015522385365329683
step: 450, loss: 8.540595445083454e-05
step: 460, loss: 0.02564260922372341
step: 470, loss: 0.00023803174553904682
step: 480, loss: 4.455312591744587e-05
step: 490, loss: 6.701423262711614e-05
step: 500, loss: 0.0011643131729215384
step: 510, loss: 8.427782449871302e-05
step: 520, loss: 0.0031007619109004736
step: 530, loss: 0.00259570125490427
/usr/local/lib/python3.7/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=0.9136822773186409, f1=0.8981097279852466,
best_f1=0.8941724941724941
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 238.67it/s]
load_f1 = 0.9173210161662818
real_f1 = 0.9176688251618872
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 246.93it/s]

```