

BDirty_50_3_5_roberta

August 11, 2022

0.1 Install packages

```
[ ]: !git clone https://github.com/megagonlabs/ditto
      %cd ditto
      !pip install -r requirements.txt
      !git clone https://github.com/rit-git/Snipnext_public
      %cd Snipnext_public
      !pip install -r requirements.txt
      %cd ..

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 15.80 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 1.2 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 51.4 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 42.6 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.3 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 65.8 MB/s
Requirement already satisfied: sklearn==0.0 in
/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
```

```

Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
|                                     | 6.4 MB 66.5 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 | 8.95 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 32.2 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
|                                     | 195 kB 71.0 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 6.9 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.3 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 60.0 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 63.2 MB/s
Requirement already satisfied: six>=1.5.0 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.15.0)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
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)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 59.8 MB/s
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: 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: 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: 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 75.4 MB/s
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
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 62.1 MB/s
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.7/dist-packages (from sklearn==0.0->-r requirements.txt
(line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)

```

```

| 880 kB 56.7 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
| 3.0 MB 22.7 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449921
sha256=7e7610757efe3fd13859b9556c35422fd12c0dc02af266e6ad5bd89510d2ca4d
  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=20ac0daec27c4ef12d9dd324837b643fb031210f74c9d6c7261b537e2a81db0b
Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
Attempting uninstall: numpy
  Found existing installation: numpy 1.21.6
  Uninstalling numpy-1.21.6:
    Successfully uninstalled numpy-1.21.6
Attempting uninstall: tqdm
  Found existing installation: tqdm 4.64.0
  Uninstalling tqdm-4.64.0:
    Successfully uninstalled tqdm-4.64.0
Attempting uninstall: srsly
  Found existing installation: srsly 2.4.4
  Uninstalling srsly-2.4.4:
    Successfully uninstalled srsly-2.4.4
Attempting uninstall: regex
  Found existing installation: regex 2022.6.2
  Uninstalling regex-2022.6.2:
    Successfully uninstalled regex-2022.6.2
Attempting uninstall: blis
  Found existing installation: blis 0.7.8
  Uninstalling blis-0.7.8:
    Successfully uninstalled blis-0.7.8
Attempting uninstall: thinc
  Found existing installation: thinc 8.1.0
  Uninstalling thinc-8.1.0:
    Successfully uninstalled thinc-8.1.0
Attempting uninstall: catalogue
  Found existing installation: catalogue 2.0.8
  Uninstalling catalogue-2.0.8:
    Successfully uninstalled catalogue-2.0.8
Attempting uninstall: torch
  Found existing installation: torch 1.12.0+cu113
  Uninstalling torch-1.12.0+cu113:
    Successfully uninstalled torch-1.12.0+cu113
Attempting uninstall: spacy
  Found existing installation: spacy 3.4.1
  Uninstalling spacy-3.4.1:
    Successfully uninstalled spacy-3.4.1
Attempting uninstall: nltk
  Found existing installation: nltk 3.7
  Uninstalling nltk-3.7:

```

```

    Successfully uninstalled nltk-3.7
Attempting uninstall: gensim
    Found existing installation: gensim 3.6.0
    Uninstalling gensim-3.6.0:
        Successfully uninstalled gensim-3.6.0
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

xarray-einstats 0.2.2 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.

torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.

torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.4.0 which is
incompatible.

torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.

tensorflow 2.8.2+zzzcolab20220719082949 requires numpy>=1.20, but you have numpy
1.19.2 which is incompatible.

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

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

en-core-web-sm 3.4.0 requires spacy<3.5.0,>=3.4.0, but you have spacy 2.2.3
which is incompatible.

cmdstanpy 1.0.4 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.
Successfully installed blis-0.4.1 catalogue-1.0.0 gensim-3.8.1 jsonlines-1.2.0
nltk-3.4.5 numpy-1.19.2 plac-1.1.3 regex-2019.12.20 sacremoses-0.0.53
sentencepiece-0.1.85 spacy-2.2.3 srsly-1.0.5 tensorboardX-2.0 thinc-7.3.1
tokenizers-0.8.1rc2 torch-1.4.0 tqdm-4.41.0 transformers-3.1.0
/content/ditto

[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.

[ ]: True

```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10085, done.
remote: Counting objects: 100% (201/201), done.
remote: Compressing objects: 100% (139/139), done.
remote: Total 10085 (delta 98), reused 130 (delta 58), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 16.66 MiB/s, done.
Resolving deltas: 100% (6903/6903), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writable
Created temporary directory: /tmp/pip-ephem-wheel-cache-y_234o8c
Created temporary directory: /tmp/pip-req-tracker-3bs9_vdi
Initialized build tracking at /tmp/pip-req-tracker-3bs9_vdi
Created build tracker: /tmp/pip-req-tracker-3bs9_vdi
Entered build tracker: /tmp/pip-req-tracker-3bs9_vdi
```

```

Created temporary directory: /tmp/pip-install-gfelonfd
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-o4xegjo_
  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-3bs9_vdi'
  Running setup.py (path:/tmp/pip-req-build-o4xegjo_/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-f4wnp641
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-f4wnp641/apex.egg-info
  writing /tmp/pip-pip-egg-info-f4wnp641/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-f4wnp641/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-f4wnp641/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-f4wnp641/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-f4wnp641/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-o4xegjo_ 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-3bs9_vdi'
Created temporary directory: /tmp/pip-unpack-js0leqyi
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-p4vxpxvt
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-p4vxpxvt
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ""'/tmp/pip-req-build-o4xegjo_/setup.py'''';

```



```
__file__='\"/tmp/pip-req-build-o4xegjo_/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-p4vxpxvt
```

```
torch.__version__ = 1.4.0
```

```
running bdist_wheel
running build
running build_py
creating build
creating build/lib
creating build/lib/apex
copying apex/__init__.py -> build/lib/apex
copying apex/_autocast_utils.py -> build/lib/apex
creating build/lib/apex/normalization
copying apex/normalization/__init__.py -> build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
creating build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
creating build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/__init__.py ->
build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/multi_tensor_apply.py ->
build/lib/apex/multi_tensor_apply
creating build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
creating build/lib/apex/parallel
```

```

    copying apex/parallel/__init__.py -> build/lib/apex/parallel
    copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
    copying apex/parallel/multiproc.py -> build/lib/apex/parallel
    copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
    copying apex/parallel/LARC.py -> build/lib/apex/parallel
    copying apex/parallel/distributed.py -> build/lib/apex/parallel
    copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
    copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
    creating build/lib/apex/optimizers
    copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
    copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
    copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
    creating build/lib/apex/fp16_utils
    copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
    creating build/lib/apex/transformer
    copying apex/transformer/microbatches.py -> build/lib/apex/transformer
    copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
    copying apex/transformer/utils.py -> build/lib/apex/transformer
    copying apex/transformer/__init__.py -> build/lib/apex/transformer
    copying apex/transformer/enums.py -> build/lib/apex/transformer
    copying apex/transformer/log_util.py -> build/lib/apex/transformer
    creating build/lib/apex/RNN
    copying apex/RNN/cells.py -> build/lib/apex/RNN
    copying apex/RNN/__init__.py -> build/lib/apex/RNN
    copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
    copying apex/RNN/models.py -> build/lib/apex/RNN
    creating build/lib/apex/contrib
    copying apex/contrib/__init__.py -> build/lib/apex/contrib
    creating build/lib/apex/amp/lists
    copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
    creating build/lib/apex/transformer/amp
    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/_data
    copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
    copying apex/transformer/_data/_batch_sampler.py ->

```

```

build/lib/apex/transformer/_data
    creating build/lib/apex/transformer/functional
    copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
    copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    creating build/lib/apex/transformer/layers
    copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
    copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
    creating build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/layers.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
    creating build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing

```

```

    copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
    creating build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py
-> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/__init__.py ->
build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/lib/apex/contrib/conv_bias_relu
    creating build/lib/apex/contrib/groupbn
    copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
    copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
    creating build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/__init__.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying
apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py ->

```

```

build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    creating build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm

```

```

    creating build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
    creating build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
    creating build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
    installing to build/bdist.linux-x86_64/wheel
    running install
    running install_lib
    creating build/bdist.linux-x86_64
    creating build/bdist.linux-x86_64/wheel
    creating build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp

```

```

    creating build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/rnn_compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/wrap.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_initialize.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_amp_state.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_process_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    creating build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    creating build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/parallel

```

```

    copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    creating build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/grad_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/microbatches.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->

```



```

build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    creating build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py
-> build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_without_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    creating build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/memory.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel

```

```

    copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/cross_entropy.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/layers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/distributed_test_base.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck_module_test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  creating build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
  copying build/lib/apex/contrib/optimizers/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers

```

```

    copying build/lib/apex/contrib/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/layer_norm.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  copying build/lib/apex/contrib/peer_memory/peer_halo_exchange_module_tests.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
  running install_egg_info
  running egg_info
  creating apex.egg-info
  writing apex.egg-info/PKG-INFO
  writing dependency_links to apex.egg-info/dependency_links.txt
  writing top-level names to apex.egg-info/top_level.txt
  writing manifest file 'apex.egg-info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file 'apex.egg-info/SOURCES.txt'
  Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
  running install_scripts
  adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
  creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
  creating '/tmp/pip-wheel-p4vxpxvt/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
  adding 'apex/__init__.py'
  adding 'apex/_autocast_utils.py'
  adding 'apex/RNN/RNNBackend.py'

```

```

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

```

```

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

```

```

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

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298114
sha256=e41ae41d3f9006d239dd0cb7bd4215aaa937c3813bff65918482db6010d54111
Stored in directory: /tmp/pip-ephem-wheel-cache-
y_234o8c/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-3bs9_vdi'
/content/ditto

```

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                               | 125 kB 16.4 MB/s
Collecting awscli
  Downloading awscli-1.25.47-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 63.6 MB/s
Collecting botocore==1.27.47
  Downloading botocore-1.27.47-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 57.5 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 66.9 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 10.7 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.47->awscli) (2.8.2)
Collecting jmespath<2.0.0,>=0.7.1
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.47->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
  Attempting uninstall: urllib3
    Found existing installation: urllib3 1.24.3
    Uninstalling urllib3-1.24.3:
      Successfully uninstalled urllib3-1.24.3
  Attempting uninstall: rsa
    Found existing installation: rsa 4.9
    Uninstalling rsa-4.9:
      Successfully uninstalled rsa-4.9
  Attempting uninstall: docutils
    Found existing installation: docutils 0.17.1
    Uninstalling docutils-0.17.1:
      Successfully uninstalled docutils-0.17.1
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

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

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

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

```

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

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
    |                                     | 776.8 MB 17 kB/s
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1) (1.19.2)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 982, done.
remote: Counting objects: 100% (52/52), done.
remote: Compressing objects: 100% (48/48), done.
remote: Total 982 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (982/982), 251.89 MiB | 17.46 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1273/1273), done.

```

0.4 remove e move

```
[ ]: !ls
```

```

apex          ditto.jpg    LICENSE     requirements.txt  run_all_wdc.py
blocking      ditto_light matcher.py  results_ditto    Snippext_public
configs.json  IC          output      run_all_er_magellan.py train_ditto.py
data          input       README.md   run_all_vary_size.py

```

```
[ ]: !rm configs.json
!mv -n /content/ditto/IC/datasasErros/BDirty_50_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 roberta \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da attr_del \
  --summarize

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

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

Downloading: 100% 481/481 [00:00<00:00, 429kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 819kB/s]

Downloading: 100% 456k/456k [00:00<00:00, 521kB/s]

Downloading: 100% 501M/501M [00:12<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: 0.45162296295166016
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=0.2828282828282828, f1=0.26262626262626265,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41407260298728943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.32558139534883723, f1=0.3132530120481928,
best_f1=0.3132530120481928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5109180808067322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3888888888888889, f1=0.26666666666666666,
best_f1=0.26666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2391260415315628
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.40740740740740744, f1=0.4680851063829786,
best_f1=0.4680851063829786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3118669092655182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4186046511627907, 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: 0.2320978045463562
/usr/local/lib/python3.7/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=0.37037037037037035, f1=0.5306122448979592,
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: 0.5700553059577942
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.588235294117647, f1=0.5555555555555556,
best_f1=0.5555555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4939366281032562
/usr/local/lib/python3.7/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=0.5116279069767441, f1=0.46511627906976755,
best_f1=0.5555555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25605908036231995
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6923076923076924, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24478654563426971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7333333333333334, f1=0.7333333333333334,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22851507365703583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.8461538461538461, f1=0.56, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09770416468381882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8461538461538461, f1=0.5384615384615384, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10851055383682251
/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.5384615384615384, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06972566246986389
/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.5833333333333334, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09487958252429962

```

```

/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.5833333333333334, best_f1=0.56
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, 119686.94it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.689655172413793
real_f1 = 0.6470588235294117
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 137.48it/s]

```

1.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
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: 0.5612180233001709
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4051995277404785
step: 20, loss: 0.5038180351257324
step: 30, loss: 0.2866060733795166
step: 40, loss: 0.3737943470478058
step: 50, loss: 0.517429769039154
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.9096483588218689
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 70, loss: 0.3610168993473053
step: 80, loss: 0.17711800336837769
step: 90, loss: 0.15583720803260803
step: 100, loss: 0.14033181965351105
step: 110, loss: 0.11287426948547363
step: 120, loss: 0.060939982533454895
step: 130, loss: 0.0048076999373734
step: 140, loss: 0.030683165416121483
step: 150, loss: 0.10988540947437286
step: 160, loss: 0.016261886805295944
step: 170, loss: 0.03907429799437523
step: 180, loss: 0.018618715927004814
step: 190, loss: 0.02967989817261696
step: 200, loss: 0.07329048961400986
step: 210, loss: 0.0060379174537956715

```

```

step: 220, loss: 0.017269005998969078
step: 230, loss: 0.0015823475550860167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9713024282560706, f1=0.9666666666666666,
best_f1=0.9666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011571052484214306
step: 10, loss: 0.12114612013101578
step: 20, loss: 0.028155213221907616
step: 30, loss: 0.017933707684278488
step: 40, loss: 0.005125515162944794
step: 50, loss: 0.0017911529866978526
step: 60, loss: 0.0024196612648665905
step: 70, loss: 0.007512195035815239
step: 80, loss: 0.008067036978900433
step: 90, loss: 0.007054508663713932
step: 100, loss: 0.001976774539798498
step: 110, loss: 0.0034429358784109354
step: 120, loss: 0.010804176330566406
step: 130, loss: 0.004967699758708477
step: 140, loss: 0.006802581250667572
step: 150, loss: 0.007142113987356424
step: 160, loss: 0.0028223686385899782
step: 170, loss: 0.0006806528544984758
step: 180, loss: 0.0025788776110857725
step: 190, loss: 0.004276310093700886
step: 200, loss: 0.002230923855677247
step: 210, loss: 0.010822968557476997
step: 220, loss: 0.0010688642505556345
step: 230, loss: 0.0006280953530222178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854423292273236, f1=0.9820627802690582,
best_f1=0.9820627802690582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0051252581179142
step: 10, loss: 0.0240080114454031
step: 20, loss: 0.020600350573658943
step: 30, loss: 0.006147139705717564
step: 40, loss: 0.011038286611437798
step: 50, loss: 0.07536715269088745
step: 60, loss: 0.0012215650640428066
step: 70, loss: 0.0010072875302284956
step: 80, loss: 0.24515880644321442
step: 90, loss: 0.10629185289144516
step: 100, loss: 0.01474425196647644
step: 110, loss: 0.006266159936785698
step: 120, loss: 0.0006340089603327215
step: 130, loss: 0.0017347830580547452
step: 140, loss: 0.032204706221818924
step: 150, loss: 0.004906738642603159
step: 160, loss: 0.003977079410105944
step: 170, loss: 0.011345073580741882
step: 180, loss: 0.03774028271436691
step: 190, loss: 0.06187891587615013
step: 200, loss: 0.020241715013980865
step: 210, loss: 0.006223498377948999
step: 220, loss: 0.03545697405934334
step: 230, loss: 0.007607212755829096
/usr/local/lib/python3.7/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=0.967670011148272, f1=0.9663677130044843,
best_f1=0.9820627802690582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010508878156542778
step: 10, loss: 0.0013683023862540722
step: 20, loss: 0.03877139091491699
step: 30, loss: 0.003187655471265316
step: 40, loss: 0.18563693761825562
step: 50, loss: 0.0054407003335654736
step: 60, loss: 0.0011161718284711242
step: 70, loss: 0.001538072363473475
step: 80, loss: 0.0005159985739737749
step: 90, loss: 0.122004933655262
step: 100, loss: 0.002461541909724474
step: 110, loss: 0.00026301603065803647
step: 120, loss: 0.0009922052267938852
step: 130, loss: 0.05890675634145737

```

```

step: 140, loss: 0.00038377143209800124
step: 150, loss: 0.022245723754167557
step: 160, loss: 0.0030603986233472824
step: 170, loss: 0.16558626294136047
step: 180, loss: 0.07730443775653839
step: 190, loss: 0.006254600360989571
step: 200, loss: 0.001367845106869936
step: 210, loss: 0.0013550639851018786
step: 220, loss: 0.0009151351405307651
step: 230, loss: 0.0028679415117949247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876819708846584, f1=0.9796839729119639,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019528917968273163
step: 10, loss: 0.00210426258854568
step: 20, loss: 0.0016102941008284688
step: 30, loss: 0.0005991054349578917
step: 40, loss: 0.006384334992617369
step: 50, loss: 0.0008837749483063817
step: 60, loss: 0.002689230488613248
step: 70, loss: 0.002100712852552533
step: 80, loss: 0.06029593572020531
step: 90, loss: 0.06163724884390831
step: 100, loss: 0.0027943237219005823
step: 110, loss: 0.0005500254919752479
step: 120, loss: 0.008865700103342533
step: 130, loss: 0.0004253305378369987
step: 140, loss: 0.0014629635261371732
step: 150, loss: 0.02774818241596222
step: 160, loss: 0.0001310136285610497
step: 170, loss: 0.028746796771883965
step: 180, loss: 0.008870154619216919
step: 190, loss: 0.04346345365047455
step: 200, loss: 0.02870067209005356
step: 210, loss: 0.0067925057373940945
step: 220, loss: 0.00010826365178218111
step: 230, loss: 0.00017364643281325698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9921259842519685, f1=0.9853768278965129,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.660729508846998e-05
step: 10, loss: 0.01231992244720459
step: 20, loss: 0.00012917495041619986
step: 30, loss: 8.947141759563237e-05
step: 40, loss: 8.106728637358174e-05
step: 50, loss: 0.00013609394954983145
step: 60, loss: 0.043225642293691635
step: 70, loss: 0.00030931076616980135
step: 80, loss: 0.006900033447891474
step: 90, loss: 0.011530122719705105
step: 100, loss: 0.0011308249086141586
step: 110, loss: 0.055587038397789
step: 120, loss: 0.00026317109586670995
step: 130, loss: 0.0016038541216403246
step: 140, loss: 0.0003926392528228462
step: 150, loss: 0.0023615090176463127
step: 160, loss: 0.0004472815780900419
step: 170, loss: 0.0003119075554423034
step: 180, loss: 0.00045121199218556285
step: 190, loss: 0.00048783706733956933
step: 200, loss: 0.007708690594881773
step: 210, loss: 0.008164787665009499
step: 220, loss: 0.0024743133690208197
step: 230, loss: 0.000982422847300768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9876819708846584, f1=0.9807909604519773,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031454653944820166
step: 10, loss: 0.0007468265830539167
step: 20, loss: 0.0002516937965992838
step: 30, loss: 0.006953253876417875
step: 40, loss: 0.0033256816677749157
step: 50, loss: 0.0016366522759199142

```



```

step: 60, loss: 0.0005572618683800101
step: 70, loss: 0.000417450355598703
step: 80, loss: 0.001369135919958353
step: 90, loss: 0.032581064850091934
step: 100, loss: 0.0008104916196316481
step: 110, loss: 0.0004629317263606936
step: 120, loss: 0.0002532207581680268
step: 130, loss: 0.00505901500582695
step: 140, loss: 0.0002641720639076084
step: 150, loss: 0.004435214214026928
step: 160, loss: 0.05704699084162712
step: 170, loss: 0.0023034075275063515
step: 180, loss: 0.00022253913630265743
step: 190, loss: 0.0005271615227684379
step: 200, loss: 0.0010990382870659232
step: 210, loss: 0.01689743809401989
step: 220, loss: 0.0003810968773905188
step: 230, loss: 0.00280338735319674
/usr/local/lib/python3.7/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=0.992108229988726, f1=0.9853107344632768,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006951221730560064
step: 10, loss: 0.0016101764049381018
step: 20, loss: 0.05055106058716774
step: 30, loss: 0.0012373252538964152
step: 40, loss: 0.0009583609644323587
step: 50, loss: 0.0031043016351759434
step: 60, loss: 0.00035765941720455885
step: 70, loss: 0.00016451324336230755
step: 80, loss: 0.04031122475862503
step: 90, loss: 0.0004579046508297324
step: 100, loss: 0.0006313806516118348
step: 110, loss: 0.000752549443859607
step: 120, loss: 0.001771995797753334
step: 130, loss: 0.0002476035733707249
step: 140, loss: 0.0001431605196557939
step: 150, loss: 0.0007418996538035572
step: 160, loss: 0.0002787442645058036
step: 170, loss: 0.013439719565212727
step: 180, loss: 9.340920951217413e-05
step: 190, loss: 0.0023482718970626593
step: 200, loss: 0.0009956619469448924
step: 210, loss: 0.0007422954658977687

```

```

step: 220, loss: 0.00026818938204087317
step: 230, loss: 0.0003168330586049706
/usr/local/lib/python3.7/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=0.990990990990991, f1=0.9853438556933484,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002517777029424906
step: 10, loss: 0.0009065814083442092
step: 20, loss: 0.0007340776501223445
step: 30, loss: 0.00038874425808899105
step: 40, loss: 0.0002746351237874478
step: 50, loss: 0.00021227862453088164
step: 60, loss: 0.00010618966916808859
step: 70, loss: 0.08546797931194305
step: 80, loss: 6.12198855378665e-05
step: 90, loss: 0.028600240126252174
step: 100, loss: 0.0062188454903662205
step: 110, loss: 9.765434515429661e-05
step: 120, loss: 0.008802189491689205
step: 130, loss: 0.0001674651721259579
step: 140, loss: 0.00015643909864593297
step: 150, loss: 0.00012641125067602843
step: 160, loss: 0.0002490603656042367
step: 170, loss: 8.30209392006509e-05
step: 180, loss: 8.145213359966874e-05
step: 190, loss: 5.063998469267972e-05
step: 200, loss: 8.001021342352033e-05
step: 210, loss: 9.085676720133051e-05
step: 220, loss: 7.962138624861836e-05
step: 230, loss: 4.4304062612354755e-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=0.9898305084745763, f1=0.9852104664391355,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.316643000696786e-05
step: 10, loss: 8.126932516461238e-05
step: 20, loss: 6.056334314052947e-05
step: 30, loss: 3.963306880905293e-05
step: 40, loss: 9.444077295484021e-05
step: 50, loss: 3.632523294072598e-05

```

```

step: 60, loss: 0.00025403406471014023
step: 70, loss: 0.0003687186981551349
step: 80, loss: 0.00011484163405839354
step: 90, loss: 8.014245395315811e-05
step: 100, loss: 7.38970484235324e-05
step: 110, loss: 9.870454960037023e-05
step: 120, loss: 5.620962838293053e-05
step: 130, loss: 0.00010827089863596484
step: 140, loss: 8.153532689902931e-05
step: 150, loss: 0.00016336205590050668
step: 160, loss: 2.6227993657812476e-05
step: 170, loss: 7.208823808468878e-05
step: 180, loss: 0.00012507168867159635
step: 190, loss: 7.58830428821966e-05
step: 200, loss: 8.678691665409133e-05
step: 210, loss: 0.0012536802096292377
step: 220, loss: 0.10703194886445999
step: 230, loss: 5.9933630836894736e-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=0.9887640449438202, f1=0.9830124575311437,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.914671808364801e-05
step: 10, loss: 0.00016119625070132315
step: 20, loss: 9.929440420819446e-05
step: 30, loss: 0.0005223123007453978
step: 40, loss: 2.3046777641866356e-05
step: 50, loss: 4.3505886424100026e-05
step: 60, loss: 0.0016036884626373649
step: 70, loss: 5.4658521548844874e-05
step: 80, loss: 9.952769323717803e-05
step: 90, loss: 0.05185144022107124
step: 100, loss: 0.00025606827693991363
step: 110, loss: 0.00014900903624948114
step: 120, loss: 0.00010391008254373446
step: 130, loss: 9.366239828523248e-05
step: 140, loss: 0.00864674523472786
step: 150, loss: 0.00020215986296534538
step: 160, loss: 0.032266270369291306
step: 170, loss: 0.00010042187932413071
step: 180, loss: 0.00011632127279881388
step: 190, loss: 0.00013624498387798667
step: 200, loss: 0.00011152312799822539
step: 210, loss: 5.177634739084169e-05

```

```

step: 220, loss: 0.0005747128743678331
step: 230, loss: 5.686139047611505e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9887387387387387, f1=0.9852774631936579,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011364607053110376
step: 10, loss: 5.168969073565677e-05
step: 20, loss: 7.404456846415997e-05
step: 30, loss: 0.002025399822741747
step: 40, loss: 5.82730644964613e-05
step: 50, loss: 0.0002607191272545606
step: 60, loss: 6.798659160267562e-05
step: 70, loss: 6.468491483246908e-05
step: 80, loss: 3.134150028927252e-05
step: 90, loss: 0.0006505451165139675
step: 100, loss: 3.663230381789617e-05
step: 110, loss: 0.0004690070345532149
step: 120, loss: 0.12350139021873474
step: 130, loss: 0.000505642790812999
step: 140, loss: 0.00021213873696979135
step: 150, loss: 0.00015329429879784584
step: 160, loss: 0.00020451376622077078
step: 170, loss: 0.00018577300943434238
step: 180, loss: 9.030280489241704e-05
step: 190, loss: 0.008241599425673485
step: 200, loss: 4.725285180029459e-05
step: 210, loss: 6.818571273470297e-05
step: 220, loss: 0.005044358782470226
step: 230, loss: 0.001345570432022214
/usr/local/lib/python3.7/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=0.9887133182844244, f1=0.9853107344632768,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.180940959486179e-05
step: 10, loss: 0.0001574481575516984
step: 20, loss: 7.061436190269887e-05
step: 30, loss: 0.0006968486122786999
step: 40, loss: 6.93818146828562e-05
step: 50, loss: 7.427832315443084e-05

```

```

step: 60, loss: 4.47125785285607e-05
step: 70, loss: 3.108518285444006e-05
step: 80, loss: 3.975598156102933e-05
step: 90, loss: 2.9577335226349533e-05
step: 100, loss: 4.8360070650232956e-05
step: 110, loss: 5.709941979148425e-05
step: 120, loss: 4.577237996272743e-05
step: 130, loss: 3.494523843983188e-05
step: 140, loss: 2.9026159609202296e-05
step: 150, loss: 0.0015935078263282776
step: 160, loss: 0.0015659657074138522
step: 170, loss: 3.3000735129462555e-05
step: 180, loss: 0.01012725941836834
step: 190, loss: 3.072111212532036e-05
step: 200, loss: 1.589526618772652e-05
step: 210, loss: 0.020584821701049805
step: 220, loss: 9.529490489512682e-05
step: 230, loss: 0.00021318436483852565
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9809203142536477,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.537109998636879e-05
step: 10, loss: 5.25072609889321e-05
step: 20, loss: 0.0004845417570322752
step: 30, loss: 0.00015784330025780946
step: 40, loss: 3.045652192668058e-05
step: 50, loss: 5.301257624523714e-05
step: 60, loss: 3.441226726863533e-05
step: 70, loss: 4.4416483433451504e-05
step: 80, loss: 3.1523748475592583e-05
step: 90, loss: 0.00011961665586568415
step: 100, loss: 3.8117206713650376e-05
step: 110, loss: 0.005152489058673382
step: 120, loss: 1.3399535419011954e-05
step: 130, loss: 0.003437751904129982
step: 140, loss: 3.43972205882892e-05
step: 150, loss: 0.023047273978590965
step: 160, loss: 0.00027017202228307724
step: 170, loss: 3.9962473238119856e-05
step: 180, loss: 2.5610466764192097e-05
step: 190, loss: 2.126692015735898e-05
step: 200, loss: 4.569940210785717e-05
step: 210, loss: 2.5703277060529217e-05

```

```

step: 220, loss: 3.268763975938782e-05
step: 230, loss: 1.8637125322129577e-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=0.9899216125419933, f1=0.9821029082774049,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005665314383804798
step: 10, loss: 2.3662340026930906e-05
step: 20, loss: 9.160328772850335e-05
step: 30, loss: 2.7137815777678043e-05
step: 40, loss: 1.9896202502422966e-05
step: 50, loss: 2.3576525563839823e-05
step: 60, loss: 0.002514332765713334
step: 70, loss: 0.00010376873979112133
step: 80, loss: 2.0887120626866817e-05
step: 90, loss: 4.3710624595405534e-05
step: 100, loss: 2.347959525650367e-05
step: 110, loss: 3.6316221667220816e-05
step: 120, loss: 5.364176831790246e-05
step: 130, loss: 2.488764039298985e-05
step: 140, loss: 2.8881166144856252e-05
step: 150, loss: 3.3224245271412656e-05
step: 160, loss: 4.272634760127403e-05
step: 170, loss: 1.5213758160825819e-05
step: 180, loss: 2.8292552087805234e-05
step: 190, loss: 5.618985233013518e-05
step: 200, loss: 3.261874007876031e-05
step: 210, loss: 0.00033941149013116956
step: 220, loss: 3.8245823816396296e-05
step: 230, loss: 3.0449178666458465e-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=0.9888392857142857, f1=0.9799107142857142,
best_f1=0.9853768278965129
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 [00:17, 144.14it/s]
load_f1 = 0.9910112359550561
real_f1 = 0.9910112359550561
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 133.97it/s]

```

1.1.3 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--da span_shuffle \
--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

```
step: 0, loss: 0.6125689744949341
```

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

```
step: 10, loss: 0.4669879078865051
```

```
step: 20, loss: 0.27135854959487915
```

```
step: 30, loss: 0.34643658995628357
```


step: 40, loss: 0.3567415773868561
step: 50, loss: 0.5203415751457214
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.32446205615997314
step: 70, loss: 0.3337625563144684
step: 80, loss: 0.32503974437713623
step: 90, loss: 0.22658535838127136
step: 100, loss: 0.3684069514274597
step: 110, loss: 0.10847997665405273
step: 120, loss: 0.14072266221046448
step: 130, loss: 0.3001860976219177
step: 140, loss: 0.24413003027439117
step: 150, loss: 0.05427974835038185
step: 160, loss: 0.14954712986946106
step: 170, loss: 0.07402832806110382
step: 180, loss: 0.13011930882930756
step: 190, loss: 0.08078975230455399
step: 200, loss: 0.15136849880218506
step: 210, loss: 0.05660811811685562
step: 220, loss: 0.0279399361461401
step: 230, loss: 0.16374759376049042
step: 240, loss: 0.07937481999397278
step: 250, loss: 0.053799454122781754
step: 260, loss: 0.20548997819423676
step: 270, loss: 0.2651089131832123
step: 280, loss: 0.04773523285984993
step: 290, loss: 0.0936405211687088
step: 300, loss: 0.2544120252132416
step: 310, loss: 0.24922409653663635
step: 320, loss: 0.1256001889705658
step: 330, loss: 0.05934574827551842
step: 340, loss: 0.5613613724708557
step: 350, loss: 0.18007385730743408
step: 360, loss: 0.06188909336924553
step: 370, loss: 0.05847114697098732
step: 380, loss: 0.15326765179634094
step: 390, loss: 0.04171428456902504
step: 400, loss: 0.05223877727985382
step: 410, loss: 0.35407400131225586
step: 420, loss: 0.050391312688589096
step: 430, loss: 0.023853030055761337
step: 440, loss: 0.02257671020925045
step: 450, loss: 0.08276189863681793
step: 460, loss: 0.053465552628040314
step: 470, loss: 0.03715104982256889
step: 480, loss: 0.1617681384086609
step: 490, loss: 0.16826000809669495
step: 500, loss: 0.04140079393982887

```

step: 510, loss: 0.04323884844779968
step: 520, loss: 0.027159802615642548
step: 530, loss: 0.10555610060691833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9178338001867413, f1=0.9190697674418604,
best_f1=0.9190697674418604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08213964849710464
step: 10, loss: 0.09296951442956924
step: 20, loss: 0.09870392829179764
step: 30, loss: 0.04708196222782135
step: 40, loss: 0.024629784747958183
step: 50, loss: 0.04472177103161812
step: 60, loss: 0.025807073339819908
step: 70, loss: 0.03891167417168617
step: 80, loss: 0.029303887858986855
step: 90, loss: 0.022124474868178368
step: 100, loss: 0.15693476796150208
step: 110, loss: 0.024266093969345093
step: 120, loss: 0.17019125819206238
step: 130, loss: 0.021737070754170418
step: 140, loss: 0.10429397225379944
step: 150, loss: 0.012828094884753227
step: 160, loss: 0.07473067194223404
step: 170, loss: 0.060986634343862534
step: 180, loss: 0.17652544379234314
step: 190, loss: 0.016240650787949562
step: 200, loss: 0.17344547808170319
step: 210, loss: 0.015467717312276363
step: 220, loss: 0.00228282087482512
step: 230, loss: 0.059743449091911316
step: 240, loss: 0.04816548526287079
step: 250, loss: 0.015594054013490677
step: 260, loss: 0.031724005937576294
step: 270, loss: 0.11378787457942963
step: 280, loss: 0.08835390210151672
step: 290, loss: 0.0471356175839901
step: 300, loss: 0.08393453806638718
step: 310, loss: 0.026123046875
step: 320, loss: 0.0478852279484272

```

```

step: 330, loss: 0.07184164971113205
step: 340, loss: 0.12417677044868469
step: 350, loss: 0.006743581499904394
step: 360, loss: 0.07517072558403015
step: 370, loss: 0.026489680632948875
step: 380, loss: 0.18511609733104706
step: 390, loss: 0.006505502853542566
step: 400, loss: 0.024242844432592392
step: 410, loss: 0.03775184974074364
step: 420, loss: 0.008700387552380562
step: 430, loss: 0.24339285492897034
step: 440, loss: 0.014934531413018703
step: 450, loss: 0.03968535736203194
step: 460, loss: 0.10330192744731903
step: 470, loss: 0.10911480337381363
step: 480, loss: 0.020779987797141075
step: 490, loss: 0.056102022528648376
step: 500, loss: 0.015110601671040058
step: 510, loss: 0.025457674637436867
step: 520, loss: 0.3779396414756775
step: 530, loss: 0.035618290305137634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9488372093023255, f1=0.9433611884865365,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1205986812710762
step: 10, loss: 0.06255210936069489
step: 20, loss: 0.00564698688685894
step: 30, loss: 0.05926861613988876
step: 40, loss: 0.028411950916051865
step: 50, loss: 0.011692274361848831
step: 60, loss: 0.016298729926347733
step: 70, loss: 0.0034006466157734394
step: 80, loss: 0.030846023932099342
step: 90, loss: 0.01529624778777361
step: 100, loss: 0.058453526347875595
step: 110, loss: 0.07944389432668686
step: 120, loss: 0.16592618823051453
step: 130, loss: 0.08541213721036911
step: 140, loss: 0.0072948685847222805

```

```

step: 150, loss: 0.07202891260385513
step: 160, loss: 0.005049470346421003
step: 170, loss: 0.0008380793151445687
step: 180, loss: 0.008930732496082783
step: 190, loss: 0.017383631318807602
step: 200, loss: 0.030343594029545784
step: 210, loss: 0.017229240387678146
step: 220, loss: 0.14455391466617584
step: 230, loss: 0.021245010197162628
step: 240, loss: 0.04745911434292793
step: 250, loss: 0.08072740584611893
step: 260, loss: 0.09324222803115845
step: 270, loss: 0.0027910657227039337
step: 280, loss: 0.00903997477144003
step: 290, loss: 0.028049632906913757
step: 300, loss: 0.081723652780056
step: 310, loss: 0.1010533943772316
step: 320, loss: 0.07180286943912506
step: 330, loss: 0.08703559637069702
step: 340, loss: 0.0061455233953893185
step: 350, loss: 0.018648212775588036
step: 360, loss: 0.01364261843264103
step: 370, loss: 0.05679170414805412
step: 380, loss: 0.028273211792111397
step: 390, loss: 0.017110368236899376
step: 400, loss: 0.11141719669103622
step: 410, loss: 0.12871204316616058
step: 420, loss: 0.0051660616882145405
step: 430, loss: 0.01907327026128769
step: 440, loss: 0.22580789029598236
step: 450, loss: 0.024552877992391586
step: 460, loss: 0.17707465589046478
step: 470, loss: 0.007299948949366808
step: 480, loss: 0.0863657146692276
step: 490, loss: 0.018133753910660744
step: 500, loss: 0.01643284037709236
step: 510, loss: 0.01093648187816143
step: 520, loss: 0.019385384395718575
step: 530, loss: 0.021013809368014336
/usr/local/lib/python3.7/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=0.9343955014058105, f1=0.9272641952135147,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02600114233791828

```

step: 10, loss: 0.003758487291634083
step: 20, loss: 0.18375082314014435
step: 30, loss: 0.004515470936894417
step: 40, loss: 0.020776690915226936
step: 50, loss: 0.10029148310422897
step: 60, loss: 0.0037240853998810053
step: 70, loss: 0.03034188784658909
step: 80, loss: 0.1487434059381485
step: 90, loss: 0.018290847539901733
step: 100, loss: 0.0014314418658614159
step: 110, loss: 0.050707872956991196
step: 120, loss: 0.0043763271532952785
step: 130, loss: 0.011097587645053864
step: 140, loss: 0.017867568880319595
step: 150, loss: 0.020618673413991928
step: 160, loss: 0.003494824515655637
step: 170, loss: 0.04450935870409012
step: 180, loss: 0.03808034956455231
step: 190, loss: 0.04553027078509331
step: 200, loss: 0.029612794518470764
step: 210, loss: 0.0006648232229053974
step: 220, loss: 0.006464425008744001
step: 230, loss: 0.0010556787019595504
step: 240, loss: 0.010929624550044537
step: 250, loss: 0.13927249610424042
step: 260, loss: 0.008657988160848618
step: 270, loss: 0.04744700714945793
step: 280, loss: 0.004907762631773949
step: 290, loss: 0.08921077102422714
step: 300, loss: 0.0024623193312436342
step: 310, loss: 0.01729525998234749
step: 320, loss: 0.14832726120948792
step: 330, loss: 0.012999010272324085
step: 340, loss: 0.0009225321118719876
step: 350, loss: 0.05088207870721817
step: 360, loss: 0.01101407315582037
step: 370, loss: 0.0029724151827394962
step: 380, loss: 0.006045909598469734
step: 390, loss: 0.0020967931486666203
step: 400, loss: 0.08368990570306778
step: 410, loss: 0.001248072017915547
step: 420, loss: 0.026555011048913002
step: 430, loss: 0.02636917307972908
step: 440, loss: 0.0019574034959077835
step: 450, loss: 0.022294772788882256
step: 460, loss: 0.012723000720143318
step: 470, loss: 0.0015359248500317335
step: 480, loss: 0.005044315941631794

```

step: 490, loss: 0.011195365339517593
step: 500, loss: 0.02183598466217518
step: 510, loss: 0.1388573795557022
step: 520, loss: 0.015420886687934399
step: 530, loss: 0.07810357213020325
/usr/local/lib/python3.7/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=0.9475138121546962, f1=0.9418764302059497,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005016646813601255
step: 10, loss: 0.004493614658713341
step: 20, loss: 0.0003245221741963178
step: 30, loss: 0.0023702492471784353
step: 40, loss: 0.0033463460858911276
step: 50, loss: 0.1486629694700241
step: 60, loss: 0.015777748078107834
step: 70, loss: 0.005706598050892353
step: 80, loss: 0.0015723510878160596
step: 90, loss: 0.08581141382455826
step: 100, loss: 0.001099717803299427
step: 110, loss: 0.0004379460879135877
step: 120, loss: 0.15899401903152466
step: 130, loss: 0.003022771328687668
step: 140, loss: 0.004738313145935535
step: 150, loss: 0.02328798919916153
step: 160, loss: 0.007760737091302872
step: 170, loss: 0.022155866026878357
step: 180, loss: 0.00787842832505703
step: 190, loss: 0.0033901366405189037
step: 200, loss: 0.02897452563047409
step: 210, loss: 0.003944251220673323
step: 220, loss: 0.0024094765540212393
step: 230, loss: 0.0026839389465749264
step: 240, loss: 0.05203806236386299
step: 250, loss: 0.195558100938797
step: 260, loss: 0.00658356212079525
step: 270, loss: 0.01447699312120676
step: 280, loss: 0.008004961535334587
step: 290, loss: 0.00318176974542439
step: 300, loss: 0.005503994878381491
step: 310, loss: 0.003634287975728512
step: 320, loss: 0.05714670941233635
step: 330, loss: 0.000569368596188724
step: 340, loss: 0.008778364397585392

```

```

step: 350, loss: 0.0004847317177336663
step: 360, loss: 0.0013627881417050958
step: 370, loss: 0.002605131594464183
step: 380, loss: 0.0016838056035339832
step: 390, loss: 0.006504877004772425
step: 400, loss: 0.003699251916259527
step: 410, loss: 0.09237466007471085
step: 420, loss: 0.19105446338653564
step: 430, loss: 0.14628364145755768
step: 440, loss: 0.005628928076475859
step: 450, loss: 0.011159846559166908
step: 460, loss: 0.01058271061629057
step: 470, loss: 0.01435413770377636
step: 480, loss: 0.027091652154922485
step: 490, loss: 0.0048516253009438515
step: 500, loss: 0.004368603695183992
step: 510, loss: 0.0045977127738296986
step: 520, loss: 0.07251294702291489
step: 530, loss: 0.016316421329975128
/usr/local/lib/python3.7/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=0.9378852536747273, f1=0.9389671361502347,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009134592488408089
step: 10, loss: 0.0007718783454038203
step: 20, loss: 0.0009160299669019878
step: 30, loss: 0.0015847109025344253
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.0010445775697007775
step: 50, loss: 0.01595889963209629
step: 60, loss: 0.005270184483379126
step: 70, loss: 0.0032866119872778654
step: 80, loss: 0.00014768984692636877
step: 90, loss: 0.00400053383782506
step: 100, loss: 0.033810943365097046
step: 110, loss: 0.017025157809257507
step: 120, loss: 0.05900915339589119
step: 130, loss: 0.0029080791864544153
step: 140, loss: 0.0017564434092491865
step: 150, loss: 0.0014110677875578403
step: 160, loss: 0.2524775266647339
step: 170, loss: 0.0013528121635317802
step: 180, loss: 0.008972210809588432
step: 190, loss: 0.04353981837630272

```

```

step: 200, loss: 0.0037702233530580997
step: 210, loss: 0.009454812854528427
step: 220, loss: 0.008271625265479088
step: 230, loss: 0.038873471319675446
step: 240, loss: 0.007956852205097675
step: 250, loss: 0.014971981756389141
step: 260, loss: 0.0005936374654993415
step: 270, loss: 0.004476822912693024
step: 280, loss: 0.014838039875030518
step: 290, loss: 0.019794873893260956
step: 300, loss: 0.0007232694188132882
step: 310, loss: 0.05459662154316902
step: 320, loss: 0.0032095147762447596
step: 330, loss: 0.002609823364764452
step: 340, loss: 0.0003943095216527581
step: 350, loss: 0.011943288147449493
step: 360, loss: 0.00038115508505143225
step: 370, loss: 0.0024802405387163162
step: 380, loss: 0.0014551185304298997
step: 390, loss: 0.00026809662813320756
step: 400, loss: 0.0016420850297436118
step: 410, loss: 0.0019265904556959867
step: 420, loss: 0.0062797050923109055
step: 430, loss: 0.0012882733717560768
step: 440, loss: 0.014404110610485077
step: 450, loss: 0.1226041167974472
step: 460, loss: 0.0064193629659712315
step: 470, loss: 0.0030296468175947666
step: 480, loss: 0.002031734911724925
step: 490, loss: 0.019686538726091385
step: 500, loss: 0.0053558447398245335
step: 510, loss: 0.12250974029302597
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 520, loss: 0.0035242103040218353
step: 530, loss: 0.007237767335027456
/usr/local/lib/python3.7/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=0.9399161620866325, f1=0.9341983317886932,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013791157398372889
step: 10, loss: 0.0012426374014467
step: 20, loss: 0.029834043234586716
step: 30, loss: 0.005992401856929064
step: 40, loss: 0.008641395717859268

```


step: 50, loss: 0.004641725681722164
step: 60, loss: 0.00102701922878623
step: 70, loss: 0.00012435462849680334
step: 80, loss: 0.0017942114500328898
step: 90, loss: 0.0038379644975066185
step: 100, loss: 0.035553254187107086
step: 110, loss: 0.0001652028295211494
step: 120, loss: 0.0008850015583448112
step: 130, loss: 0.0003973552957177162
step: 140, loss: 0.002764371922239661
step: 150, loss: 0.0028443941846489906
step: 160, loss: 0.002944862935692072
step: 170, loss: 0.0006729881861247122
step: 180, loss: 0.00048692847485654056
step: 190, loss: 0.050019387155771255
step: 200, loss: 0.041668884456157684
step: 210, loss: 0.00019227006123401225
step: 220, loss: 0.0010867989622056484
step: 230, loss: 0.00040412324597127736
step: 240, loss: 0.014973815530538559
step: 250, loss: 0.08577238023281097
step: 260, loss: 0.0020380758214741945
step: 270, loss: 0.003492548828944564
step: 280, loss: 0.0022196262143552303
step: 290, loss: 0.006623462308198214
step: 300, loss: 0.001416421844623983
step: 310, loss: 0.00044579297536984086
step: 320, loss: 0.009193210862576962
step: 330, loss: 0.0006837264518253505
step: 340, loss: 0.00014754744188394397
step: 350, loss: 0.00025573206949047744
step: 360, loss: 0.0037453756667673588
step: 370, loss: 0.10041017830371857
step: 380, loss: 0.08895343542098999
step: 390, loss: 0.0006543683120980859
step: 400, loss: 0.15795449912548065
step: 410, loss: 0.01225123181939125
step: 420, loss: 0.016147617250680923
step: 430, loss: 0.001822198973968625
step: 440, loss: 0.00789332203567028
step: 450, loss: 0.010167286731302738
step: 460, loss: 0.00015319099475163966
step: 470, loss: 0.11929583549499512
step: 480, loss: 0.006224066484719515
step: 490, loss: 0.00042169983498752117
step: 500, loss: 0.012183820828795433
step: 510, loss: 0.0007125558331608772
step: 520, loss: 0.003294797847047448

```
step: 530, loss: 0.0015422505093738437
/usr/local/lib/python3.7/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=0.9419953596287702, f1=0.937442502299908,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006281697424128652
step: 10, loss: 0.0006192477885633707
step: 20, loss: 0.00015872875519562513
step: 30, loss: 0.00012584975047502667
step: 40, loss: 3.916233981726691e-05
step: 50, loss: 0.0020128206815570593
step: 60, loss: 0.00010498215851839632
step: 70, loss: 0.0024507432244718075
step: 80, loss: 0.0002865121641661972
step: 90, loss: 0.0014142772415652871
step: 100, loss: 0.00025609374279156327
step: 110, loss: 0.0006067213253118098
step: 120, loss: 0.00032666517654433846
step: 130, loss: 0.021905528381466866
step: 140, loss: 0.022158408537507057
step: 150, loss: 0.0024381547700613737
step: 160, loss: 0.05601887032389641
step: 170, loss: 0.020967094227671623
step: 180, loss: 0.0028349817730486393
step: 190, loss: 0.031087325885891914
step: 200, loss: 0.0010929031996056437
step: 210, loss: 0.10683687031269073
step: 220, loss: 0.0004914657911285758
step: 230, loss: 0.01132653746753931
step: 240, loss: 0.0029632195364683867
step: 250, loss: 0.0038039549253880978
step: 260, loss: 0.0010769553482532501
step: 270, loss: 0.00244908407330513
step: 280, loss: 0.0002444216806907207
step: 290, loss: 0.0017686433857306838
step: 300, loss: 0.0001843819918576628
step: 310, loss: 0.001327571109868586
step: 320, loss: 0.00019249484466854483
step: 330, loss: 0.007179189007729292
step: 340, loss: 0.004248879849910736
step: 350, loss: 0.002794337924569845
step: 360, loss: 0.005524324718862772
step: 370, loss: 0.06478334963321686
step: 380, loss: 0.00010208233288722113
```

```

step: 390, loss: 0.004336387850344181
step: 400, loss: 0.0007730096112936735
step: 410, loss: 0.005036166403442621
step: 420, loss: 0.0022856032010167837
step: 430, loss: 0.0012496714480221272
step: 440, loss: 0.02811262756586075
step: 450, loss: 0.0005050746840424836
step: 460, loss: 0.013948556035757065
step: 470, loss: 0.04525274038314819
step: 480, loss: 0.007723681628704071
step: 490, loss: 0.0008230910170823336
step: 500, loss: 0.00019249068282078952
step: 510, loss: 0.008235152810811996
step: 520, loss: 0.002860818523913622
step: 530, loss: 0.00039768466376699507
/usr/local/lib/python3.7/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=0.9417698303530491, f1=0.9339366515837104,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00034141266951337457
step: 10, loss: 0.0037769994232803583
step: 20, loss: 6.519193266285583e-05
step: 30, loss: 0.004567537922412157
step: 40, loss: 0.00012680186773650348
step: 50, loss: 0.0001737329293973744
step: 60, loss: 0.0002949604531750083
step: 70, loss: 0.0008062497363425791
step: 80, loss: 0.004545173142105341
step: 90, loss: 0.014043115079402924
step: 100, loss: 0.0007765524787828326
step: 110, loss: 0.02487843856215477
step: 120, loss: 6.482437311206013e-05
step: 130, loss: 0.0003080478636547923
step: 140, loss: 0.00011998770060017705
step: 150, loss: 0.0004278332053218037
step: 160, loss: 0.0496116578578949
step: 170, loss: 0.007815289311110973
step: 180, loss: 0.01954875886440277
step: 190, loss: 0.0003879238211084157
step: 200, loss: 0.0009890658548101783
step: 210, loss: 0.00030403875280171633
step: 220, loss: 0.0008050727774389088
step: 230, loss: 0.00039216617005877197
step: 240, loss: 0.00013281672727316618

```

```

step: 250, loss: 0.00018002960132434964
step: 260, loss: 0.0024026131723076105
step: 270, loss: 0.009793421253561974
step: 280, loss: 0.016467882320284843
step: 290, loss: 4.576165520120412e-05
step: 300, loss: 0.05095900222659111
step: 310, loss: 0.0022145649418234825
step: 320, loss: 0.00023356068413704634
step: 330, loss: 0.00012963962217327207
step: 340, loss: 0.037447284907102585
step: 350, loss: 0.010181600227952003
step: 360, loss: 0.0006092709954828024
step: 370, loss: 0.0010531465522944927
step: 380, loss: 0.0007352230604737997
step: 390, loss: 0.001404010457918048
step: 400, loss: 0.022307412698864937
step: 410, loss: 0.00023703464830759913
step: 420, loss: 0.004884404595941305
step: 430, loss: 0.0024354110937565565
step: 440, loss: 5.792256706627086e-05
step: 450, loss: 0.0018171739066019654
step: 460, loss: 0.0007744573522359133
step: 470, loss: 8.96135315997526e-05
step: 480, loss: 0.000149925981531851
step: 490, loss: 0.0006800710689276457
step: 500, loss: 2.3848628188716248e-05
step: 510, loss: 0.0001830199034884572
step: 520, loss: 4.324032488511875e-05
step: 530, loss: 0.00022528215777128935
/usr/local/lib/python3.7/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=0.9468331021729081, f1=0.9390187987161852,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034665821585804224
step: 10, loss: 6.274820043472573e-05
step: 20, loss: 8.297328895423561e-05
step: 30, loss: 0.021724961698055267
step: 40, loss: 0.0003850608190987259
step: 50, loss: 0.004290980286896229
step: 60, loss: 0.031857024878263474
step: 70, loss: 2.027999835263472e-05
step: 80, loss: 1.418189458490815e-05
step: 90, loss: 0.0006119219469837844
step: 100, loss: 0.01718122512102127

```

```
step: 110, loss: 0.005191898439079523
step: 120, loss: 1.3526177099265624e-05
step: 130, loss: 0.012878982350230217
step: 140, loss: 0.05125419422984123
step: 150, loss: 3.180397106916644e-05
step: 160, loss: 0.07228098809719086
step: 170, loss: 0.0002285251539433375
step: 180, loss: 0.049796782433986664
step: 190, loss: 5.765902460552752e-05
step: 200, loss: 0.040111660957336426
step: 210, loss: 0.01882888749241829
step: 220, loss: 0.00027084138127975166
step: 230, loss: 0.0002374516479903832
step: 240, loss: 3.2870590075617656e-05
step: 250, loss: 3.1141822546487674e-05
step: 260, loss: 0.0018971929093822837
step: 270, loss: 3.0780698580201715e-05
step: 280, loss: 0.012894725427031517
step: 290, loss: 5.834121839143336e-05
step: 300, loss: 5.477112063090317e-05
step: 310, loss: 0.042252473533153534
step: 320, loss: 0.00044280063593760133
step: 330, loss: 0.00011989978520432487
step: 340, loss: 1.8949856894323602e-05
step: 350, loss: 0.00010333002137485892
step: 360, loss: 5.551276262849569e-05
step: 370, loss: 0.0033330644946545362
step: 380, loss: 0.0007360040326602757
step: 390, loss: 6.504698103526607e-05
step: 400, loss: 0.0003411909274291247
step: 410, loss: 0.00013940116332378238
step: 420, loss: 4.346412970335223e-05
step: 430, loss: 3.610543717513792e-05
step: 440, loss: 3.9803635445423424e-05
step: 450, loss: 0.04699265584349632
step: 460, loss: 7.840068428777158e-05
step: 470, loss: 0.011752482503652573
step: 480, loss: 0.0001094106410164386
step: 490, loss: 0.0003257718635722995
step: 500, loss: 8.401463855989277e-05
step: 510, loss: 3.870572618325241e-05
step: 520, loss: 0.001009360421448946
step: 530, loss: 0.003949485253542662
/usr/local/lib/python3.7/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=0.9443166129774505, f1=0.9391304347826086,
best_f1=0.9433611884865365
```

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

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

step: 0, loss: 0.00038274165126495063
step: 10, loss: 0.00017425538680981845
step: 20, loss: 0.0004181403783150017
step: 30, loss: 0.000854659010656178
step: 40, loss: 0.0005245045176707208
step: 50, loss: 2.358042729611043e-05
step: 60, loss: 3.662043309304863e-05
step: 70, loss: 3.6623470805352554e-05
step: 80, loss: 0.00012455506657715887
step: 90, loss: 0.0038923569954931736
step: 100, loss: 4.85098353237845e-05
step: 110, loss: 2.5394214389962144e-05
step: 120, loss: 0.00037792877992615104
step: 130, loss: 2.4112330720527098e-05
step: 140, loss: 2.1080810256535187e-05
step: 150, loss: 0.00021848225151188672
step: 160, loss: 0.0001971907913684845
step: 170, loss: 0.006922692526131868
step: 180, loss: 1.7154581655631773e-05
step: 190, loss: 0.0019016911974176764
step: 200, loss: 2.966536339954473e-05
step: 210, loss: 0.00047357703442685306
step: 220, loss: 0.0042282408103346825
step: 230, loss: 8.638891813461669e-06
step: 240, loss: 0.00025811532395891845
step: 250, loss: 0.000725550577044487
step: 260, loss: 1.5910376532701775e-05
step: 270, loss: 4.281816291040741e-05
step: 280, loss: 0.17966508865356445
step: 290, loss: 0.0003756918595172465
step: 300, loss: 1.1537085811141878e-05
step: 310, loss: 0.0003812720242422074
step: 320, loss: 6.0006859712302685e-05
step: 330, loss: 2.5720877601997927e-05
step: 340, loss: 6.0062149714212865e-05
step: 350, loss: 0.00023637183767277747
step: 360, loss: 5.132502701599151e-05
step: 370, loss: 0.0004358699661679566
step: 380, loss: 0.006980427540838718
step: 390, loss: 0.000224981369683519
step: 400, loss: 0.004365968983620405
step: 410, loss: 0.03583629056811333
step: 420, loss: 4.830293983104639e-05
step: 430, loss: 0.00044685413013212383
step: 440, loss: 0.00016235903603956103

```
step: 450, loss: 0.00019484994118101895
step: 460, loss: 6.14336458966136e-05
step: 470, loss: 0.0014787119580432773
step: 480, loss: 0.000510800164192915
step: 490, loss: 0.00011265240027569234
step: 500, loss: 3.793080759351142e-05
step: 510, loss: 0.010065593756735325
step: 520, loss: 8.55631151353009e-05
step: 530, loss: 0.05559000000357628
/usr/local/lib/python3.7/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=0.9415554532903819, f1=0.9368131868131868,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009081271477043629
step: 10, loss: 0.0017123341094702482
step: 20, loss: 0.001188294612802565
step: 30, loss: 1.1361890756234061e-05
step: 40, loss: 0.0007163009140640497
step: 50, loss: 1.366027572657913e-05
step: 60, loss: 1.5791240002727136e-05
step: 70, loss: 0.00021739580552093685
step: 80, loss: 0.001265243743546307
step: 90, loss: 0.004255768842995167
step: 100, loss: 0.0006081528845243156
step: 110, loss: 2.8443721021176316e-05
step: 120, loss: 3.1277118978323415e-05
step: 130, loss: 5.08636912854854e-05
step: 140, loss: 0.00036578145227395
step: 150, loss: 8.388002606807277e-05
step: 160, loss: 0.002212801016867161
step: 170, loss: 1.5362566045951098e-05
step: 180, loss: 0.008484147489070892
step: 190, loss: 1.4833937711955514e-05
step: 200, loss: 0.001398091553710401
step: 210, loss: 0.006987974513322115
step: 220, loss: 1.2773826711054426e-05
step: 230, loss: 4.7780806198716164e-05
step: 240, loss: 9.69020402408205e-05
step: 250, loss: 1.4811464097874705e-05
step: 260, loss: 0.0021813546773046255
step: 270, loss: 0.005236657336354256
step: 280, loss: 2.8798393032047898e-05
step: 290, loss: 0.0003173043660353869
step: 300, loss: 0.00017592246877029538
```

```

step: 310, loss: 0.0002181710151489824
step: 320, loss: 0.00012822636927012354
step: 330, loss: 0.012159841135144234
step: 340, loss: 0.0002932821516878903
step: 350, loss: 1.0840380127774552e-05
step: 360, loss: 2.1315197955118492e-05
step: 370, loss: 0.00017803370428737253
step: 380, loss: 3.037306305486709e-05
step: 390, loss: 0.0013095998438075185
step: 400, loss: 3.7798203265992925e-05
step: 410, loss: 0.00048490779590792954
step: 420, loss: 0.0002420122327748686
step: 430, loss: 0.0001560645323479548
step: 440, loss: 0.002273187506943941
step: 450, loss: 0.0011111756321042776
step: 460, loss: 1.848754436650779e-05
step: 470, loss: 0.001792466384358704
step: 480, loss: 0.00011780697241192684
step: 490, loss: 2.9853423257009126e-05
step: 500, loss: 1.9179467926733196e-05
step: 510, loss: 5.803786189062521e-05
step: 520, loss: 0.0024466116447001696
step: 530, loss: 2.9881392038078047e-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=0.9474654377880185, f1=0.9408014571948998,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027563574258238077
step: 10, loss: 0.00028034186107106507
step: 20, loss: 1.844714643084444e-05
step: 30, loss: 9.003942068375181e-06
step: 40, loss: 0.0020841420628130436
step: 50, loss: 0.0006211043801158667
step: 60, loss: 3.144503716612235e-05
step: 70, loss: 0.0009731233585625887
step: 80, loss: 0.013124549761414528
step: 90, loss: 2.5874142011161894e-05
step: 100, loss: 0.02631889097392559
step: 110, loss: 4.93269253638573e-05
step: 120, loss: 0.00027125756605528295
step: 130, loss: 9.170051635010168e-05
step: 140, loss: 4.298035491956398e-05
step: 150, loss: 1.3164940355636645e-05
step: 160, loss: 0.0003273098554927856

```



```

step: 170, loss: 0.00015137431910261512
step: 180, loss: 0.00015543415793217719
step: 190, loss: 7.925165846245363e-05
step: 200, loss: 6.24030944891274e-05
step: 210, loss: 8.765498932916671e-06
step: 220, loss: 0.00013439827307593077
step: 230, loss: 1.4967919923947193e-05
step: 240, loss: 5.3384854254545644e-05
step: 250, loss: 6.245093391044065e-05
step: 260, loss: 2.1460002244566567e-05
step: 270, loss: 0.0009604772785678506
step: 280, loss: 1.1611580703174695e-05
step: 290, loss: 3.200651553925127e-05
step: 300, loss: 0.0001521070080343634
step: 310, loss: 3.991398625657894e-05
step: 320, loss: 3.134943108307198e-05
step: 330, loss: 1.5955143680912443e-05
step: 340, loss: 1.799244637368247e-05
step: 350, loss: 1.0844120879482944e-05
step: 360, loss: 0.00014604900206904858
step: 370, loss: 0.0009324154816567898
step: 380, loss: 1.0195975846727379e-05
step: 390, loss: 6.092137118685059e-05
step: 400, loss: 0.007235875353217125
step: 410, loss: 3.054921035072766e-05
step: 420, loss: 6.308084266493097e-05
step: 430, loss: 1.7504389688838273e-05
step: 440, loss: 6.9364482442324515e-06
step: 450, loss: 1.3585780834546313e-05
step: 460, loss: 1.9258692191215232e-05
step: 470, loss: 0.0010794592089951038
step: 480, loss: 5.234014224697603e-06
step: 490, loss: 1.5406589227495715e-05
step: 500, loss: 1.835559851315338e-05
step: 510, loss: 9.439687346457504e-06
step: 520, loss: 5.3623076382791623e-05
step: 530, loss: 7.290335361176403e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9443665264142123, f1=0.9400369003690037,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.410308868798893e-05
step: 10, loss: 0.0001487109693698585
step: 20, loss: 8.296125997730996e-06

```

step: 30, loss: 0.0009577342425473034
step: 40, loss: 1.875104135251604e-05
step: 50, loss: 1.447212343919091e-05
step: 60, loss: 3.1253664928954095e-05
step: 70, loss: 1.4654744518338703e-05
step: 80, loss: 8.676132893015165e-06
step: 90, loss: 5.651240826409776e-06
step: 100, loss: 1.4080970686336514e-05
step: 110, loss: 0.00017817705520428717
step: 120, loss: 6.2249318943941034e-06
step: 130, loss: 8.754357622819953e-06
step: 140, loss: 1.4964027286623605e-05
step: 150, loss: 1.0520031537453178e-05
step: 160, loss: 0.0003686040290631354
step: 170, loss: 1.0240461051580496e-05
step: 180, loss: 2.0891429812763818e-05
step: 190, loss: 6.72781970934011e-06
step: 200, loss: 9.417434739589226e-06
step: 210, loss: 4.10149987146724e-05
step: 220, loss: 6.247282271942822e-06
step: 230, loss: 0.000273895391728729
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 240, loss: 9.867992957879324e-06
step: 250, loss: 0.017048271372914314
step: 260, loss: 6.828414370829705e-06
step: 270, loss: 5.923294520471245e-05
step: 280, loss: 3.195303361280821e-05
step: 290, loss: 5.215387773205293e-06
step: 300, loss: 6.82096833770629e-06
step: 310, loss: 3.2726049539633095e-05
step: 320, loss: 6.191405191202648e-06
step: 330, loss: 2.5856330466922373e-05
step: 340, loss: 2.1384745195973665e-05
step: 350, loss: 7.342496701312484e-06
step: 360, loss: 0.11464536190032959
step: 370, loss: 8.165749022737145e-06
step: 380, loss: 0.0013084693346172571
step: 390, loss: 5.1222555157437455e-06
step: 400, loss: 0.0004360082093626261
step: 410, loss: 6.5341182562406175e-06
step: 420, loss: 3.43763422279153e-05
step: 430, loss: 0.0006708605214953423
step: 440, loss: 3.469386138021946e-05
step: 450, loss: 0.0005449820891954005
step: 460, loss: 5.408073411672376e-05
step: 470, loss: 6.891747943882365e-06
step: 480, loss: 2.291546479682438e-05
step: 490, loss: 6.735285751346964e-06

```

step: 500, loss: 0.001407892326824367
step: 510, loss: 0.002195150824263692
step: 520, loss: 7.413232196995523e-06
step: 530, loss: 0.00041754008270800114
/usr/local/lib/python3.7/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=0.947075208913649, f1=0.9412304866850323,
best_f1=0.9433611884865365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.319695446814876e-06
step: 10, loss: 2.384267281740904e-05
step: 20, loss: 3.8142701669130474e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0002441492397338152
step: 40, loss: 7.793240911269095e-06
step: 50, loss: 8.497278031427413e-06
step: 60, loss: 3.124677459709346e-05
step: 70, loss: 9.793480785447173e-06
step: 80, loss: 5.5432101362384856e-06
step: 90, loss: 1.1346674000378698e-05
step: 100, loss: 4.1313355723104905e-06
step: 110, loss: 0.010681267827749252
step: 120, loss: 1.2780907127307728e-05
step: 130, loss: 4.128261934965849e-05
step: 140, loss: 6.57506143397768e-06
step: 150, loss: 5.800255166832358e-06
step: 160, loss: 0.0002728800172917545
step: 170, loss: 4.760902811540291e-06
step: 180, loss: 1.0330083568987902e-05
step: 190, loss: 1.808124034141656e-05
step: 200, loss: 3.309881140012294e-05
step: 210, loss: 7.189748430391774e-06
step: 220, loss: 5.967889137536986e-06
step: 230, loss: 3.874909452861175e-05
step: 240, loss: 5.714573035220383e-06
step: 250, loss: 5.092452738608699e-06
step: 260, loss: 5.513405994861387e-06
step: 270, loss: 1.8374543287791312e-05
step: 280, loss: 9.09538211999461e-05
step: 290, loss: 2.5647806978668086e-05
step: 300, loss: 0.017085859552025795
step: 310, loss: 0.05795358866453171
step: 320, loss: 7.662863936275244e-06
step: 330, loss: 6.768823823222192e-06
step: 340, loss: 7.133893177524442e-06

```

```

step: 350, loss: 0.010731488466262817
step: 360, loss: 7.253084277181188e-06
step: 370, loss: 2.534709346946329e-05
step: 380, loss: 1.937633533088956e-05
step: 390, loss: 8.106150744424667e-06
step: 400, loss: 0.02793271839618683
step: 410, loss: 6.906651378812967e-06
step: 420, loss: 7.759723303024657e-06
step: 430, loss: 5.803979092888767e-06
step: 440, loss: 5.64378933631815e-06
step: 450, loss: 4.807096047443338e-05
step: 460, loss: 1.2944846275786404e-05
step: 470, loss: 6.563931037817383e-06
step: 480, loss: 8.270067155535799e-06
step: 490, loss: 6.779991053917911e-06
step: 500, loss: 5.278718617773848e-06
step: 510, loss: 4.4839111069450155e-05
step: 520, loss: 2.1718158677686006e-05
step: 530, loss: 5.763001354353037e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9463869463869464, f1=0.9388505747126438,
best_f1=0.9433611884865365
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 [00:34, 167.38it/s]
load_f1 = 0.9494199535962878
real_f1 = 0.9460966542750929
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 143.67it/s]

```

1.1.4 Amazon-Google - Running the matcher

```

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

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

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

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

```
step: 0, loss: 0.5143054127693176
```

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

```
step: 10, loss: 0.39790982007980347
```

```
step: 20, loss: 0.45242124795913696
```

```
step: 30, loss: 0.3294215500354767
```

```
step: 40, loss: 0.3230559825897217
```

```
step: 50, loss: 0.4809192717075348
```

```
step: 60, loss: 0.49341222643852234
```

```
step: 70, loss: 0.30111873149871826
```

```
step: 80, loss: 0.3786582946777344
```

```
step: 90, loss: 0.23746979236602783
```

```
step: 100, loss: 0.2350785732269287
```

```

step: 110, loss: 0.24347138404846191
step: 120, loss: 0.37710046768188477
step: 130, loss: 0.2209327667951584
step: 140, loss: 0.5137990117073059
step: 150, loss: 0.3157199025154114
step: 160, loss: 0.4936700463294983
step: 170, loss: 0.205021932721138
step: 180, loss: 0.2625451982021332
step: 190, loss: 0.3969932198524475
step: 200, loss: 0.21178603172302246
step: 210, loss: 0.5301807522773743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4280701754385965, f1=0.467680608365019,
best_f1=0.467680608365019
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37088659405708313
step: 10, loss: 0.04939355328679085
step: 20, loss: 0.3871482014656067
step: 30, loss: 0.2772270739078522
step: 40, loss: 0.388569176197052
step: 50, loss: 0.23856763541698456
step: 60, loss: 0.3290488123893738
step: 70, loss: 0.2874446511268616
step: 80, loss: 0.1895352154970169
step: 90, loss: 0.15929219126701355
step: 100, loss: 0.6325422525405884
step: 110, loss: 0.3213280439376831
step: 120, loss: 0.12345539033412933
step: 130, loss: 0.1563139408826828
step: 140, loss: 0.24062660336494446
step: 150, loss: 0.23632720112800598
step: 160, loss: 0.0940321758389473
step: 170, loss: 0.18552108108997345
step: 180, loss: 0.2797599136829376
step: 190, loss: 0.31730663776397705
step: 200, loss: 0.09911829233169556
step: 210, loss: 0.149664044380188
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5350553505535055, f1=0.5576923076923076,
best_f1=0.5576923076923076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06865700334310532
step: 10, loss: 0.12479791045188904
step: 20, loss: 0.30360648036003113
step: 30, loss: 0.10176768153905869
step: 40, loss: 0.2674962282180786
step: 50, loss: 0.3157825171947479
step: 60, loss: 0.14568282663822174
step: 70, loss: 0.18651342391967773
step: 80, loss: 0.25323694944381714
step: 90, loss: 0.09549771994352341
step: 100, loss: 0.19071052968502045
step: 110, loss: 0.08799533545970917
step: 120, loss: 0.26921606063842773
step: 130, loss: 0.20978926122188568
step: 140, loss: 0.23150864243507385
step: 150, loss: 0.20611758530139923
step: 160, loss: 0.3102477788925171
step: 170, loss: 0.23265686631202698
step: 180, loss: 0.15763971209526062
step: 190, loss: 0.018585745245218277
step: 200, loss: 0.14801447093486786
step: 210, loss: 0.19678594172000885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6142595978062158, f1=0.6242774566473989,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13086600601673126
step: 10, loss: 0.13064470887184143
step: 20, loss: 0.050351519137620926
step: 30, loss: 0.02279926836490631
step: 40, loss: 0.059776026755571365

```



```

step: 50, loss: 0.10817908495664597
step: 60, loss: 0.2669784724712372
step: 70, loss: 0.11714082956314087
step: 80, loss: 0.13325762748718262
step: 90, loss: 0.07571430504322052
step: 100, loss: 0.16061143577098846
step: 110, loss: 0.5555009841918945
step: 120, loss: 0.17720593512058258
step: 130, loss: 0.1932452917098999
step: 140, loss: 0.4715109169483185
step: 150, loss: 0.03435535356402397
step: 160, loss: 0.23047560453414917
step: 170, loss: 0.13950002193450928
step: 180, loss: 0.021542109549045563
step: 190, loss: 0.15384447574615479
step: 200, loss: 0.10152534395456314
step: 210, loss: 0.3936578333377838
/usr/local/lib/python3.7/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=0.577391304347826, f1=0.5888689407540395,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21497514843940735
step: 10, loss: 0.13630515336990356
step: 20, loss: 0.027074936777353287
step: 30, loss: 0.012127207592129707
step: 40, loss: 0.0452125146985054
step: 50, loss: 0.24796423316001892
step: 60, loss: 0.21865582466125488
step: 70, loss: 0.15032029151916504
step: 80, loss: 0.11616434901952744
step: 90, loss: 0.09620513767004013
step: 100, loss: 0.0721825659275055
step: 110, loss: 0.04274654760956764
step: 120, loss: 0.1522299349308014
step: 130, loss: 0.08713321387767792
step: 140, loss: 0.35323643684387207
step: 150, loss: 0.05511251837015152
step: 160, loss: 0.04357067495584488
step: 170, loss: 0.023637939244508743
step: 180, loss: 0.15972529351711273
step: 190, loss: 0.1308324635028839
step: 200, loss: 0.31567442417144775
step: 210, loss: 0.10020168870687485
/usr/local/lib/python3.7/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=0.598890942698706, f1=0.588235294117647,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026550596579909325
step: 10, loss: 0.24996720254421234
step: 20, loss: 0.04941905289888382
step: 30, loss: 0.12283452600240707
step: 40, loss: 0.09254702180624008
step: 50, loss: 0.0496886670589447
step: 60, loss: 0.18429258465766907
step: 70, loss: 0.09151620417833328
step: 80, loss: 0.056830182671546936
step: 90, loss: 0.14376385509967804
step: 100, loss: 0.05837484449148178
step: 110, loss: 0.023777201771736145
step: 120, loss: 0.02527947537600994
step: 130, loss: 0.08395133912563324
step: 140, loss: 0.048479631543159485
step: 150, loss: 0.1472841501235962
step: 160, loss: 0.045813724398612976
step: 170, loss: 0.06222740188241005
step: 180, loss: 0.1352767050266266
step: 190, loss: 0.0867934450507164
step: 200, loss: 0.17748117446899414
step: 210, loss: 0.039197683334350586
/usr/local/lib/python3.7/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=0.6055045871559633, f1=0.5759637188208616,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08287747949361801
step: 10, loss: 0.1253892481327057
step: 20, loss: 0.020022079348564148
step: 30, loss: 0.08899209648370743
step: 40, loss: 0.015691326931118965
step: 50, loss: 0.09683073312044144
step: 60, loss: 0.030715694651007652
step: 70, loss: 0.016191275790333748
step: 80, loss: 0.03059752658009529
step: 90, loss: 0.11319691687822342
step: 100, loss: 0.11731728166341782

```

```

step: 110, loss: 0.2656742334365845
step: 120, loss: 0.12683069705963135
step: 130, loss: 0.2187711000442505
step: 140, loss: 0.05979710444808006
step: 150, loss: 0.03865897282958031
step: 160, loss: 0.2712622284889221
step: 170, loss: 0.12398522347211838
step: 180, loss: 0.01335006020963192
step: 190, loss: 0.16400498151779175
step: 200, loss: 0.1950911581516266
step: 210, loss: 0.08966581523418427
/usr/local/lib/python3.7/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=0.5944881889763781, f1=0.6016260162601625,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0414896123111248
step: 10, loss: 0.07711833715438843
step: 20, loss: 0.007145251613110304
step: 30, loss: 0.11722081899642944
step: 40, loss: 0.1768091768026352
step: 50, loss: 0.16288724541664124
step: 60, loss: 0.060213975608348846
step: 70, loss: 0.17453385889530182
step: 80, loss: 0.059568852186203
step: 90, loss: 0.13655956089496613
step: 100, loss: 0.14310602843761444
step: 110, loss: 0.055212363600730896
step: 120, loss: 0.25230497121810913
step: 130, loss: 0.03795204311609268
step: 140, loss: 0.018778998404741287
step: 150, loss: 0.09789497405290604
step: 160, loss: 0.06659701466560364
step: 170, loss: 0.059094831347465515
step: 180, loss: 0.04806368798017502
step: 190, loss: 0.0860789343714714
step: 200, loss: 0.0777391791343689
step: 210, loss: 0.0669906809926033
/usr/local/lib/python3.7/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=0.598890942698706, f1=0.6086956521739131,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.054386936128139496
step: 10, loss: 0.22487890720367432
step: 20, loss: 0.07625081390142441
step: 30, loss: 0.009341521188616753
step: 40, loss: 0.04730619490146637
step: 50, loss: 0.13459435105323792
step: 60, loss: 0.14285780489444733
step: 70, loss: 0.2943016588687897
step: 80, loss: 0.023022059351205826
step: 90, loss: 0.01467632781714201
step: 100, loss: 0.011845136061310768
step: 110, loss: 0.07530001550912857
step: 120, loss: 0.019958555698394775
step: 130, loss: 0.08929996937513351
step: 140, loss: 0.07796591520309448
step: 150, loss: 0.004435409791767597
step: 160, loss: 0.0753864049911499
step: 170, loss: 0.0122873829677701
step: 180, loss: 0.05693718418478966
step: 190, loss: 0.07524874806404114
step: 200, loss: 0.014450340531766415
step: 210, loss: 0.04513696953654289
/usr/local/lib/python3.7/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=0.5919117647058825, f1=0.5908256880733944,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019364302279427648
step: 10, loss: 0.1280401647090912
step: 20, loss: 0.12819741666316986
step: 30, loss: 0.0015498597640544176
step: 40, loss: 0.04923643171787262
step: 50, loss: 0.05372818559408188
step: 60, loss: 0.005730170290917158
step: 70, loss: 0.042924586683511734
step: 80, loss: 0.024509992450475693
step: 90, loss: 0.011148352175951004
step: 100, loss: 0.0855950191617012
step: 110, loss: 0.00156163249630481
step: 120, loss: 0.01971103996038437
step: 130, loss: 0.0027434329967945814
step: 140, loss: 0.04657655954360962
step: 150, loss: 0.08215915411710739
step: 160, loss: 0.052576709538698196

```

```

step: 170, loss: 0.013081824406981468
step: 180, loss: 0.012695797719061375
step: 190, loss: 0.03766622766852379
step: 200, loss: 0.2962260842323303
step: 210, loss: 0.028691761195659637
/usr/local/lib/python3.7/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=0.606822262118492, f1=0.594306049822064,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05327015742659569
step: 10, loss: 0.03971415013074875
step: 20, loss: 0.033821746706962585
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.04082215949892998
step: 40, loss: 0.007386954501271248
step: 50, loss: 0.01793358474969864
step: 60, loss: 0.09178868681192398
step: 70, loss: 0.0647081583738327
step: 80, loss: 0.18974265456199646
step: 90, loss: 0.04731064662337303
step: 100, loss: 0.12394161522388458
step: 110, loss: 0.08165514469146729
step: 120, loss: 0.03189362585544586
step: 130, loss: 0.02049778774380684
step: 140, loss: 0.04117847979068756
step: 150, loss: 0.02283734828233719
step: 160, loss: 0.006342655513435602
step: 170, loss: 0.011529923416674137
step: 180, loss: 0.004080403596162796
step: 190, loss: 0.1163949966430664
step: 200, loss: 0.012995068915188313
step: 210, loss: 0.16148556768894196
/usr/local/lib/python3.7/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=0.5980582524271845, f1=0.6080305927342257,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011407018639147282
step: 10, loss: 0.00784939993172884
step: 20, loss: 0.10153613239526749
step: 30, loss: 0.04263880476355553

```

```

step: 40, loss: 0.0017386997351422906
step: 50, loss: 0.0017474349588155746
step: 60, loss: 0.0005070849438197911
step: 70, loss: 0.0040831128135323524
step: 80, loss: 0.0972922071814537
step: 90, loss: 0.022725477814674377
step: 100, loss: 0.004012099467217922
step: 110, loss: 0.18164141476154327
step: 120, loss: 0.003622426651418209
step: 130, loss: 0.1827719360589981
step: 140, loss: 0.00850613135844469
step: 150, loss: 0.0013325916370376945
step: 160, loss: 0.010867512784898281
step: 170, loss: 0.011329138651490211
step: 180, loss: 0.0074854358099401
step: 190, loss: 0.014906533993780613
step: 200, loss: 0.0035423727240413427
step: 210, loss: 0.013725976459681988
/usr/local/lib/python3.7/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=0.6012024048096193, f1=0.6047430830039526,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009281705133616924
step: 10, loss: 0.006387779954820871
step: 20, loss: 0.05702817067503929
step: 30, loss: 0.012527775950729847
step: 40, loss: 0.05992317199707031
step: 50, loss: 0.11631006747484207
step: 60, loss: 0.037317708134651184
step: 70, loss: 0.006306702271103859
step: 80, loss: 0.011217452585697174
step: 90, loss: 0.23687055706977844
step: 100, loss: 0.023189125582575798
step: 110, loss: 0.013897187076508999
step: 120, loss: 0.011402860283851624
step: 130, loss: 0.014423206448554993
step: 140, loss: 0.009100155904889107
step: 150, loss: 0.0016629230231046677
step: 160, loss: 0.06267593801021576
step: 170, loss: 0.12615075707435608
step: 180, loss: 0.01483957376331091
step: 190, loss: 0.058613989502191544
step: 200, loss: 0.04795670136809349
step: 210, loss: 0.004469193983823061

```

```

/usr/local/lib/python3.7/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=0.5928853754940712, f1=0.6204238921001927,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003120481502264738
step: 10, loss: 0.015865888446569443
step: 20, loss: 0.020561423152685165
step: 30, loss: 0.0030586523935198784
step: 40, loss: 0.021216467022895813
step: 50, loss: 0.004522922448813915
step: 60, loss: 0.2057344913482666
step: 70, loss: 0.002555761719122529
step: 80, loss: 0.09826165437698364
step: 90, loss: 0.002611541422083974
step: 100, loss: 0.010001085698604584
step: 110, loss: 0.0006663817912340164
step: 120, loss: 0.002548373769968748
step: 130, loss: 0.018110085278749466
step: 140, loss: 0.0025975017342716455
step: 150, loss: 0.035603445023298264
step: 160, loss: 0.052841901779174805
step: 170, loss: 0.059485722333192825
step: 180, loss: 0.0031983479857444763
step: 190, loss: 0.06946957856416702
step: 200, loss: 0.0024044315796345472
step: 210, loss: 0.0018815497169271111
/usr/local/lib/python3.7/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=0.5903614457831325, f1=0.6011787819253438,
best_f1=0.6242774566473989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021086575463414192
step: 10, loss: 0.007419477216899395
step: 20, loss: 0.03858331963419914
step: 30, loss: 0.0026427782140672207
step: 40, loss: 0.01968887448310852
step: 50, loss: 0.008663933724164963
step: 60, loss: 0.04734093323349953
step: 70, loss: 0.0010330918012186885
step: 80, loss: 0.0069253710098564625
step: 90, loss: 0.0015581774059683084

```

```

step: 100, loss: 0.02055763639509678
step: 110, loss: 0.011697297915816307
step: 120, loss: 0.008209534920752048
step: 130, loss: 0.0038256009574979544
step: 140, loss: 0.008567546494305134
step: 150, loss: 0.03270217403769493
step: 160, loss: 0.01578652486205101
step: 170, loss: 0.02091319113969803
step: 180, loss: 0.001530247274786234
step: 190, loss: 0.002354827942326665
step: 200, loss: 0.005218786187469959
step: 210, loss: 0.052490051835775375
/usr/local/lib/python3.7/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=0.6003976143141153, f1=0.6076923076923078,
best_f1=0.6242774566473989
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 [00:10, 227.58it/s]

```



```
load_f1 = 0.6033519553072626
real_f1 = 0.590818363273453
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 [00:31, 141.62it/s]
```

1.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

```
enabled          : True
opt_level        : O2
```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 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: 0.4669240713119507
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.36991196870803833
step: 20, loss: 0.2777514159679413
step: 30, loss: 0.3748180568218231
step: 40, loss: 0.2348354309797287
step: 50, loss: 0.302741140127182
step: 60, loss: 0.5453515648841858
step: 70, loss: 0.4501717686653137
step: 80, loss: 0.15167056024074554
step: 90, loss: 0.29209667444229126
step: 100, loss: 0.44203415513038635
step: 110, loss: 0.23675481975078583
step: 120, loss: 0.3229730725288391
step: 130, loss: 0.32026049494743347
step: 140, loss: 0.16696405410766602
step: 150, loss: 0.33096006512641907
step: 160, loss: 0.23713819682598114

```

```

step: 170, loss: 0.37792062759399414
step: 180, loss: 0.15927444398403168
step: 190, loss: 0.15804320573806763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3854984939098358
step: 10, loss: 0.33113184571266174
step: 20, loss: 0.6187273859977722
step: 30, loss: 0.2372918725013733
step: 40, loss: 0.5619180202484131
step: 50, loss: 0.31472980976104736
step: 60, loss: 0.46063461899757385
step: 70, loss: 0.2982519865036011
step: 80, loss: 0.147679403424263
step: 90, loss: 0.302202969789505
step: 100, loss: 0.24167975783348083
step: 110, loss: 0.3853992521762848
step: 120, loss: 0.24603138864040375
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.4805518686771393
step: 140, loss: 0.30359944701194763
step: 150, loss: 0.3030119836330414
step: 160, loss: 0.3188723921775818
step: 170, loss: 0.23319146037101746
step: 180, loss: 0.17464548349380493
step: 190, loss: 0.23536312580108643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2016, f1=0.20474137931034486, best_f1=0.20474137931034486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3622778654098511

```

```

step: 10, loss: 0.3733825981616974
step: 20, loss: 0.42747431993484497
step: 30, loss: 0.2950683534145355
step: 40, loss: 0.09284865856170654
step: 50, loss: 0.3677239716053009
step: 60, loss: 0.15648937225341797
step: 70, loss: 0.37344688177108765
step: 80, loss: 0.2663692533969879
step: 90, loss: 0.3794202506542206
step: 100, loss: 0.5336235165596008
step: 110, loss: 0.6695621609687805
step: 120, loss: 0.3973022699356079
step: 130, loss: 0.1502295881509781
step: 140, loss: 0.3653688430786133
step: 150, loss: 0.31584176421165466
step: 160, loss: 0.5541137456893921
step: 170, loss: 0.38994067907333374
step: 180, loss: 0.3441615402698517
step: 190, loss: 0.1698993295431137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2039911308203991, f1=0.20646625070901875,
best_f1=0.20646625070901875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2255609780550003
step: 10, loss: 0.20483337342739105
step: 20, loss: 0.2689249515533447
step: 30, loss: 0.2274949848651886
step: 40, loss: 0.5251474976539612
step: 50, loss: 0.2244097888469696
step: 60, loss: 0.33802080154418945
step: 70, loss: 0.2666705846786499
step: 80, loss: 0.2515817880630493
step: 90, loss: 0.16825547814369202
step: 100, loss: 0.3016875386238098
step: 110, loss: 0.42061540484428406
step: 120, loss: 0.22863279283046722
step: 130, loss: 0.4107550382614136
step: 140, loss: 0.2845378518104553
step: 150, loss: 0.21166610717773438
step: 160, loss: 0.26196375489234924

```

```

step: 170, loss: 0.47093361616134644
step: 180, loss: 0.34810250997543335
step: 190, loss: 0.14812473952770233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.21930400525279056, f1=0.21908602150537634,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36920323967933655
step: 10, loss: 0.35555118322372437
step: 20, loss: 0.15149368345737457
step: 30, loss: 0.13164715468883514
step: 40, loss: 0.3180857002735138
step: 50, loss: 0.4783186614513397
step: 60, loss: 0.22891750931739807
step: 70, loss: 0.35076218843460083
step: 80, loss: 0.3352087736129761
step: 90, loss: 0.28253015875816345
step: 100, loss: 0.40679600834846497
step: 110, loss: 0.40826478600502014
step: 120, loss: 0.21881861984729767
step: 130, loss: 0.5314061045646667
step: 140, loss: 0.3456679582595825
step: 150, loss: 0.27178940176963806
step: 160, loss: 0.14809755980968475
step: 170, loss: 0.32513388991355896
step: 180, loss: 0.22404387593269348
step: 190, loss: 0.26548174023628235
/usr/local/lib/python3.7/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=0.21849865951742628, f1=0.2165867032213845,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.258816659450531
step: 10, loss: 0.2948940694332123
step: 20, loss: 0.2750505805015564
step: 30, loss: 0.4155055582523346
step: 40, loss: 0.2668996751308441

```

```

step: 50, loss: 0.2856453061103821
step: 60, loss: 0.41677191853523254
step: 70, loss: 0.3217248022556305
step: 80, loss: 0.2837084233760834
step: 90, loss: 0.2080969661474228
step: 100, loss: 0.4278374910354614
step: 110, loss: 0.21560774743556976
step: 120, loss: 0.4205699563026428
step: 130, loss: 0.5468884706497192
step: 140, loss: 0.17090871930122375
step: 150, loss: 0.42556071281433105
step: 160, loss: 0.3803744614124298
step: 170, loss: 0.37396040558815
step: 180, loss: 0.15634912252426147
step: 190, loss: 0.2804371118545532
/usr/local/lib/python3.7/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=0.21691678035470668, f1=0.21273617914625612,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2921214699745178
step: 10, loss: 0.26714444160461426
step: 20, loss: 0.35234421491622925
step: 30, loss: 0.20427635312080383
step: 40, loss: 0.29292991757392883
step: 50, loss: 0.08029893040657043
step: 60, loss: 0.1452973335981369
step: 70, loss: 0.15823277831077576
step: 80, loss: 0.22604651749134064
step: 90, loss: 0.2291349619626999
step: 100, loss: 0.5468398928642273
step: 110, loss: 0.41538041830062866
step: 120, loss: 0.35490846633911133
step: 130, loss: 0.279849112033844
step: 140, loss: 0.22012194991111755
step: 150, loss: 0.27841782569885254
step: 160, loss: 0.3598898649215698
step: 170, loss: 0.27290716767311096
step: 180, loss: 0.2829156816005707
step: 190, loss: 0.28698691725730896
/usr/local/lib/python3.7/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=0.2190237797246558, f1=0.21621621621621623,
best_f1=0.21908602150537634

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21132192015647888
step: 10, loss: 0.35740384459495544
step: 20, loss: 0.22069402039051056
step: 30, loss: 0.34281468391418457
step: 40, loss: 0.1489936113357544
step: 50, loss: 0.29197534918785095
step: 60, loss: 0.5787938833236694
step: 70, loss: 0.22529791295528412
step: 80, loss: 0.38043755292892456
step: 90, loss: 0.23227526247501373
step: 100, loss: 0.285869300365448
step: 110, loss: 0.3489311635494232
step: 120, loss: 0.3477364182472229
step: 130, loss: 0.2835324704647064
step: 140, loss: 0.3512634038925171
step: 150, loss: 0.2934792637825012
step: 160, loss: 0.1777947098016739
step: 170, loss: 0.23163476586341858
step: 180, loss: 0.2890676259994507
step: 190, loss: 0.2708495259284973
/usr/local/lib/python3.7/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=0.20293000542593595, f1=0.20515633571036754,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23640356957912445
step: 10, loss: 0.09308461844921112
step: 20, loss: 0.29786449670791626
step: 30, loss: 0.17511482536792755
step: 40, loss: 0.295089989900589
step: 50, loss: 0.3544245958328247
step: 60, loss: 0.36889418959617615
step: 70, loss: 0.1587175875902176
step: 80, loss: 0.3674421012401581
step: 90, loss: 0.6455711722373962
step: 100, loss: 0.3470747768878937
step: 110, loss: 0.3394218683242798
step: 120, loss: 0.5522647500038147
step: 130, loss: 0.2907246947288513
step: 140, loss: 0.2788575291633606
step: 150, loss: 0.233114093542099
step: 160, loss: 0.21446382999420166

```

```

step: 170, loss: 0.5353246927261353
step: 180, loss: 0.4708310067653656
step: 190, loss: 0.23655445873737335
/usr/local/lib/python3.7/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=0.2026286966046002, f1=0.2076200993926008,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24232929944992065
step: 10, loss: 0.1547691971063614
step: 20, loss: 0.29231882095336914
step: 30, loss: 0.3580997884273529
step: 40, loss: 0.292770117521286
step: 50, loss: 0.0954504981637001
step: 60, loss: 0.309684693813324
step: 70, loss: 0.2814731001853943
step: 80, loss: 0.3664720952510834
step: 90, loss: 0.1504446417093277
step: 100, loss: 0.21470053493976593
step: 110, loss: 0.22398996353149414
step: 120, loss: 0.3494841158390045
step: 130, loss: 0.45369842648506165
step: 140, loss: 0.3336455523967743
step: 150, loss: 0.287922739982605
step: 160, loss: 0.22766050696372986
step: 170, loss: 0.33373239636421204
step: 180, loss: 0.20708687603473663
step: 190, loss: 0.1519087702035904
/usr/local/lib/python3.7/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=0.20353982300884957, f1=0.20835654596100275,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5123540163040161
step: 10, loss: 0.08137404918670654
step: 20, loss: 0.2319105863571167
step: 30, loss: 0.2254839837551117
step: 40, loss: 0.07830998301506042
step: 50, loss: 0.2840941250324249
step: 60, loss: 0.22962240874767303
step: 70, loss: 0.5577206015586853
step: 80, loss: 0.46096935868263245

```



```

step: 90, loss: 0.3830150067806244
step: 100, loss: 0.16276654601097107
step: 110, loss: 0.2344798892736435
step: 120, loss: 0.2790588438510895
step: 130, loss: 0.5751437544822693
step: 140, loss: 0.42418715357780457
step: 150, loss: 0.2978120446205139
step: 160, loss: 0.3127763271331787
step: 170, loss: 0.41126760840415955
step: 180, loss: 0.2841591536998749
step: 190, loss: 0.09908012300729752
/usr/local/lib/python3.7/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=0.2026578073089701, f1=0.2089385474860335,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3608098328113556
step: 10, loss: 0.280115008354187
step: 20, loss: 0.49039313197135925
step: 30, loss: 0.3611364960670471
step: 40, loss: 0.2877599596977234
step: 50, loss: 0.48567330837249756
step: 60, loss: 0.4206307530403137
step: 70, loss: 0.35740870237350464
step: 80, loss: 0.3664638102054596
step: 90, loss: 0.41633138060569763
step: 100, loss: 0.2903006374835968
step: 110, loss: 0.5079002976417542
step: 120, loss: 0.1690879911184311
step: 130, loss: 0.22506895661354065
step: 140, loss: 0.37006351351737976
step: 150, loss: 0.4292570948600769
step: 160, loss: 0.22044916450977325
step: 170, loss: 0.4181603491306305
step: 180, loss: 0.28417620062828064
step: 190, loss: 0.15547794103622437
/usr/local/lib/python3.7/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=0.20322043309272625, f1=0.209106239460371,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2891145646572113

```

```

step: 10, loss: 0.3541569709777832
step: 20, loss: 0.3687742054462433
step: 30, loss: 0.3687148988246918
step: 40, loss: 0.21592219173908234
step: 50, loss: 0.2184790074825287
step: 60, loss: 0.28760960698127747
step: 70, loss: 0.1592084914445877
step: 80, loss: 0.29949161410331726
step: 90, loss: 0.2925685942173004
step: 100, loss: 0.16759341955184937
step: 110, loss: 0.55184006690979
step: 120, loss: 0.1478452980518341
step: 130, loss: 0.22111225128173828
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.2908417284488678
step: 150, loss: 0.36381715536117554
step: 160, loss: 0.40342405438423157
step: 170, loss: 0.11217425018548965
step: 180, loss: 0.23362185060977936
step: 190, loss: 0.30490830540657043
/usr/local/lib/python3.7/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=0.2026578073089701, f1=0.2089385474860335,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28724467754364014
step: 10, loss: 0.21562306582927704
step: 20, loss: 0.3471706211566925
step: 30, loss: 0.11676312983036041
step: 40, loss: 0.2867278754711151
step: 50, loss: 0.3569433391094208
step: 60, loss: 0.2781628370285034
step: 70, loss: 0.2338952124118805
step: 80, loss: 0.42593351006507874
step: 90, loss: 0.1431102752685547
step: 100, loss: 0.29071444272994995
step: 110, loss: 0.41634294390678406
step: 120, loss: 0.279325008392334
step: 130, loss: 0.3035580813884735
step: 140, loss: 0.15693853795528412
step: 150, loss: 0.42130228877067566
step: 160, loss: 0.3620651662349701
step: 170, loss: 0.2263888567686081
step: 180, loss: 0.30905529856681824
step: 190, loss: 0.5584570169448853

```

```

/usr/local/lib/python3.7/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=0.20365246264526837, f1=0.208821887213847,
best_f1=0.21908602150537634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5773481726646423
step: 10, loss: 0.2175232470035553
step: 20, loss: 0.2353806495666504
step: 30, loss: 0.29502585530281067
step: 40, loss: 0.6119107604026794
step: 50, loss: 0.29290929436683655
step: 60, loss: 0.16779592633247375
step: 70, loss: 0.23459607362747192
step: 80, loss: 0.20718616247177124
step: 90, loss: 0.3044244050979614
step: 100, loss: 0.34606388211250305
step: 110, loss: 0.4239329993724823
step: 120, loss: 0.22126281261444092
step: 130, loss: 0.2913080155849457
step: 140, loss: 0.5027002692222595
step: 150, loss: 0.15910197794437408
step: 160, loss: 0.27627918124198914
step: 170, loss: 0.7451610565185547
step: 180, loss: 0.2920967638492584
step: 190, loss: 0.4127427041530609
/usr/local/lib/python3.7/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=0.20365246264526837, f1=0.2089385474860335,
best_f1=0.21908602150537634
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 [00:13, 154.79it/s]
load_f1 = 0.22002635046113309
real_f1 = 0.21872953503601833
733it [00:00, 3351.87it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 145.25it/s]

```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

```

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

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

```

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

```

UserWarning)
step: 0, loss: 0.49576640129089355
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5019739270210266
step: 20, loss: 0.30468419194221497
step: 30, loss: 0.4030097424983978
step: 40, loss: 0.481125146150589
step: 50, loss: 0.3219566345214844
step: 60, loss: 0.5195575952529907
step: 70, loss: 0.35865315794944763
step: 80, loss: 0.29291361570358276
step: 90, loss: 0.2200697362422943
step: 100, loss: 0.14563260972499847
step: 110, loss: 0.3967827260494232
step: 120, loss: 0.31539368629455566
step: 130, loss: 0.30191659927368164
step: 140, loss: 0.3881320655345917
step: 150, loss: 0.3159852921962738
step: 160, loss: 0.39205604791641235
step: 170, loss: 0.32748809456825256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.1967054263565891, f1=0.19512195121951217,
best_f1=0.19512195121951217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3103208541870117
step: 10, loss: 0.4720926582813263
step: 20, loss: 0.32651636004447937
step: 30, loss: 0.3002035617828369
step: 40, loss: 0.09517794847488403
step: 50, loss: 0.41464996337890625
step: 60, loss: 0.17938227951526642
step: 70, loss: 0.48711106181144714
step: 80, loss: 0.22582906484603882
step: 90, loss: 0.22775481641292572
step: 100, loss: 0.5018126964569092
step: 110, loss: 0.255844384431839
step: 120, loss: 0.23086902499198914
step: 130, loss: 0.5326627492904663
step: 140, loss: 0.4913075566291809
step: 150, loss: 0.43634268641471863

```

```

step: 160, loss: 0.43028098344802856
step: 170, loss: 0.36647000908851624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24624060150375937, f1=0.23363544813695866,
best_f1=0.23363544813695866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6286676526069641
step: 10, loss: 0.32185685634613037
step: 20, loss: 0.2129788100719452
step: 30, loss: 0.235867440700531
step: 40, loss: 0.3433595299720764
step: 50, loss: 0.5818670988082886
step: 60, loss: 0.2735922932624817
step: 70, loss: 0.2619507908821106
step: 80, loss: 0.40716683864593506
step: 90, loss: 0.489837646484375
step: 100, loss: 0.2454843521118164
step: 110, loss: 0.1861116588115692
step: 120, loss: 0.5043148398399353
step: 130, loss: 0.45387348532676697
step: 140, loss: 0.43055522441864014
step: 150, loss: 0.20441479980945587
step: 160, loss: 0.16238108277320862
step: 170, loss: 0.30076736211776733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2550218340611354, f1=0.24867724867724864,
best_f1=0.24867724867724864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40631306171417236
step: 10, loss: 0.48373526334762573
step: 20, loss: 0.21748247742652893
step: 30, loss: 0.3665342628955841

```

```

step: 40, loss: 0.2443804144859314
step: 50, loss: 0.3457317054271698
step: 60, loss: 0.6638603210449219
step: 70, loss: 0.32052990794181824
step: 80, loss: 0.45445016026496887
step: 90, loss: 0.27942749857902527
step: 100, loss: 0.38135093450546265
step: 110, loss: 0.4136575162410736
step: 120, loss: 0.4534135162830353
step: 130, loss: 0.3174439072608948
step: 140, loss: 0.22617235779762268
step: 150, loss: 0.6928300261497498
step: 160, loss: 0.14117780327796936
step: 170, loss: 0.24939438700675964
/usr/local/lib/python3.7/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=0.25274725274725274, f1=0.24880382775119617,
best_f1=0.24867724867724864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3480154573917389
step: 10, loss: 0.2974579930305481
step: 20, loss: 0.2764416038990021
step: 30, loss: 0.23686861991882324
step: 40, loss: 0.2354438304901123
step: 50, loss: 0.2407056838274002
step: 60, loss: 0.293486624956131
step: 70, loss: 0.34985867142677307
step: 80, loss: 0.08013135194778442
step: 90, loss: 0.5876641869544983
step: 100, loss: 0.24778442084789276
step: 110, loss: 0.26578208804130554
step: 120, loss: 0.11932899802923203
step: 130, loss: 0.22320352494716644
step: 140, loss: 0.20889244973659515
step: 150, loss: 0.30128172039985657
step: 160, loss: 0.2638166546821594
step: 170, loss: 0.35918375849723816
/usr/local/lib/python3.7/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=0.2537313432835821, f1=0.25378787878787884,
best_f1=0.24867724867724864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.29237276315689087
step: 10, loss: 0.12320098280906677
step: 20, loss: 0.31112217903137207
step: 30, loss: 0.2270042896270752
step: 40, loss: 0.4200607240200043
step: 50, loss: 0.22542302310466766
step: 60, loss: 0.36240002512931824
step: 70, loss: 0.4083130657672882
step: 80, loss: 0.19999606907367706
step: 90, loss: 0.28267228603363037
step: 100, loss: 0.15675237774848938
step: 110, loss: 0.40444934368133545
step: 120, loss: 0.3888418972492218
step: 130, loss: 0.5461939573287964
step: 140, loss: 0.2595749795436859
step: 150, loss: 0.18376164138317108
step: 160, loss: 0.3175453245639801
step: 170, loss: 0.22860148549079895
/usr/local/lib/python3.7/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=0.2518518518518519, f1=0.2522522522522523,
best_f1=0.24867724867724864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22538593411445618
step: 10, loss: 0.465462327003479
step: 20, loss: 0.41853630542755127
step: 30, loss: 0.45578885078430176
step: 40, loss: 0.12279614061117172
step: 50, loss: 0.37930551171302795
step: 60, loss: 0.4657808840274811
step: 70, loss: 0.31603649258613586
step: 80, loss: 0.15530867874622345
step: 90, loss: 0.48234835267066956
step: 100, loss: 0.26589930057525635
step: 110, loss: 0.32979559898376465
step: 120, loss: 0.3350425660610199
step: 130, loss: 0.19038541615009308
step: 140, loss: 0.15125690400600433
step: 150, loss: 0.25251853466033936
step: 160, loss: 0.31499457359313965
step: 170, loss: 0.24673905968666077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.5564738292011019, f1=0.5318559556786704,

best_f1=0.5318559556786704

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

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

step: 0, loss: 0.445562481880188

step: 10, loss: 0.3791588246822357

step: 20, loss: 0.21392260491847992

step: 30, loss: 0.11422274261713028

step: 40, loss: 0.09131139516830444

step: 50, loss: 0.07401686161756516

step: 60, loss: 0.2000846266746521

step: 70, loss: 0.17632974684238434

step: 80, loss: 0.22901299595832825

step: 90, loss: 0.10524815320968628

step: 100, loss: 0.06723771244287491

step: 110, loss: 0.22643564641475677

step: 120, loss: 0.3265344202518463

step: 130, loss: 0.12044670432806015

step: 140, loss: 0.2866784632205963

step: 150, loss: 0.05501466244459152

step: 160, loss: 0.03778986260294914

step: 170, loss: 0.23809461295604706

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

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

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

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=0.7584415584415585, f1=0.7908163265306123,

best_f1=0.7908163265306123

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

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

step: 0, loss: 0.2576202154159546

step: 10, loss: 0.06702855974435806

step: 20, loss: 0.03182979300618172

step: 30, loss: 0.06981118023395538

step: 40, loss: 0.4147201180458069

step: 50, loss: 0.024240391328930855

step: 60, loss: 0.15640206634998322

step: 70, loss: 0.2351672351360321

step: 80, loss: 0.056743573397397995

step: 90, loss: 0.049652453511953354

```

step: 100, loss: 0.1553463339805603
step: 110, loss: 0.16335336863994598
step: 120, loss: 0.2625492811203003
step: 130, loss: 0.030421704053878784
step: 140, loss: 0.06084121763706207
step: 150, loss: 0.37883177399635315
step: 160, loss: 0.012270476669073105
step: 170, loss: 0.10198166966438293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7884615384615384, f1=0.8367816091954023,
best_f1=0.8367816091954023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020113985985517502
step: 10, loss: 0.0727020874619484
step: 20, loss: 0.1374526470899582
step: 30, loss: 0.1821410208940506
step: 40, loss: 0.11559254676103592
step: 50, loss: 0.06375443935394287
step: 60, loss: 0.15364181995391846
step: 70, loss: 0.03995687887072563
step: 80, loss: 0.07543274760246277
step: 90, loss: 0.0187862329185009
step: 100, loss: 0.07897569239139557
step: 110, loss: 0.022914880886673927
step: 120, loss: 0.033721666783094406
step: 130, loss: 0.13040491938591003
step: 140, loss: 0.04455816373229027
step: 150, loss: 0.050679054111242294
step: 160, loss: 0.006037788465619087
step: 170, loss: 0.19569909572601318
/usr/local/lib/python3.7/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=0.7692307692307693, f1=0.8190954773869347,
best_f1=0.8367816091954023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16415947675704956
step: 10, loss: 0.02348964661359787

```

```

step: 20, loss: 0.01248883455991745
step: 30, loss: 0.04088974744081497
step: 40, loss: 0.0054058292880654335
step: 50, loss: 0.098362497985363
step: 60, loss: 0.04008973762392998
step: 70, loss: 0.023579446598887444
step: 80, loss: 0.08383704721927643
step: 90, loss: 0.038044825196266174
step: 100, loss: 0.02696395106613636
step: 110, loss: 0.02724323607981205
step: 120, loss: 0.08550344407558441
step: 130, loss: 0.02480163425207138
step: 140, loss: 0.24200133979320526
step: 150, loss: 0.03321990743279457
step: 160, loss: 0.1703348159790039
step: 170, loss: 0.07733714580535889
/usr/local/lib/python3.7/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=0.7880434782608696, f1=0.8329048843187661,
best_f1=0.8367816091954023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02936921827495098
step: 10, loss: 0.08265910297632217
step: 20, loss: 0.016645468771457672
step: 30, loss: 0.0906796008348465
step: 40, loss: 0.00492830527946353
step: 50, loss: 0.010838009417057037
step: 60, loss: 0.1401170939207077
step: 70, loss: 0.010480214841663837
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.018256776034832
step: 90, loss: 0.05467343330383301
step: 100, loss: 0.013538378290832043
step: 110, loss: 0.04912896826863289
step: 120, loss: 0.03811896592378616
step: 130, loss: 0.1346028596162796
step: 140, loss: 0.04209892079234123
step: 150, loss: 0.01723472774028778
step: 160, loss: 0.17090682685375214
step: 170, loss: 0.23671017587184906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7851851851851852, f1=0.8254716981132074,
best_f1=0.8367816091954023

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03542185574769974
step: 10, loss: 0.004636527504771948
step: 20, loss: 0.023523176088929176
step: 30, loss: 0.04295377433300018
step: 40, loss: 0.005206012632697821
step: 50, loss: 0.060396332293748856
step: 60, loss: 0.011002767831087112
step: 70, loss: 0.15641666948795319
step: 80, loss: 0.0029229053761810064
step: 90, loss: 0.07263290137052536
step: 100, loss: 0.024516064673662186
step: 110, loss: 0.008288197219371796
step: 120, loss: 0.009212218225002289
step: 130, loss: 0.10452943295240402
step: 140, loss: 0.014998833648860455
step: 150, loss: 0.03557879105210304
step: 160, loss: 0.1567763090133667
step: 170, loss: 0.005892253015190363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.8052631578947368, f1=0.8454106280193237,
best_f1=0.8454106280193237
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14615902304649353
step: 10, loss: 0.027803480625152588
step: 20, loss: 0.07210379093885422
step: 30, loss: 0.08142627030611038
step: 40, loss: 0.007018347270786762
step: 50, loss: 0.005583673249930143
step: 60, loss: 0.07846260070800781
step: 70, loss: 0.014305965974926949
step: 80, loss: 0.009691277518868446
step: 90, loss: 0.0024875355884432793
step: 100, loss: 0.024617115035653114
step: 110, loss: 0.012972935102880001
step: 120, loss: 0.09132754802703857
step: 130, loss: 0.09817066788673401
step: 140, loss: 0.053030919283628464

```

```

step: 150, loss: 0.029454907402396202
step: 160, loss: 0.006640914361923933
step: 170, loss: 0.04958358407020569
/usr/local/lib/python3.7/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=0.7967479674796748, f1=0.846938775510204,
best_f1=0.8454106280193237
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001744783134199679
step: 10, loss: 0.08768606930971146
step: 20, loss: 0.0031302704010158777
step: 30, loss: 0.13942554593086243
step: 40, loss: 0.027276216074824333
step: 50, loss: 0.06891338527202606
step: 60, loss: 0.03184037655591965
step: 70, loss: 0.08127789199352264
step: 80, loss: 0.031185533851385117
step: 90, loss: 0.014118642546236515
step: 100, loss: 0.06045340374112129
step: 110, loss: 0.03511900454759598
step: 120, loss: 0.004330833908170462
step: 130, loss: 0.04185478389263153
step: 140, loss: 0.011239073239266872
step: 150, loss: 0.0035740716848522425
step: 160, loss: 0.00459966529160738
step: 170, loss: 0.018078450113534927
/usr/local/lib/python3.7/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=0.7872340425531914, f1=0.8450000000000001,
best_f1=0.8454106280193237
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 225.62it/s]
load_f1 = 0.5017667844522967
real_f1 = 0.44610778443113774
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 141.73it/s]

```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

```

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

```

```

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

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 562kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 7.57MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 4.82MB/s]
Downloading: 100% 501M/501M [00:18<00:00, 27.7MB/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: 0.5643205642700195

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

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

step: 10, loss: 0.43370378017425537

step: 20, loss: 0.5365296602249146

step: 30, loss: 0.31280168890953064

step: 40, loss: 0.32641279697418213

step: 50, loss: 0.6023504734039307

step: 60, loss: 0.46203500032424927

step: 70, loss: 0.38507160544395447

step: 80, loss: 0.22044765949249268

step: 90, loss: 0.2635113298892975

step: 100, loss: 0.2237797975540161

step: 110, loss: 0.20792432129383087

step: 120, loss: 0.07852010428905487

step: 130, loss: 0.12182540446519852

step: 140, loss: 0.11669807136058807

step: 150, loss: 0.26680588722229004

step: 160, loss: 0.030499376356601715

step: 170, loss: 0.17673394083976746

step: 180, loss: 0.11776050925254822

step: 190, loss: 0.0411110557615757

step: 200, loss: 0.0440615713596344

step: 210, loss: 0.16354694962501526

step: 220, loss: 0.05807969719171524

step: 230, loss: 0.0034275357611477375

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

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

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

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=0.9522727272727274, f1=0.9462857142857144,

best_f1=0.9462857142857144

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

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

step: 0, loss: 0.010266709141433239

step: 10, loss: 0.1355915367603302

step: 20, loss: 0.014622186310589314

step: 30, loss: 0.044188275933265686

```

step: 40, loss: 0.048647165298461914
step: 50, loss: 0.00628884369507432
step: 60, loss: 0.009748700074851513
step: 70, loss: 0.01952977478504181
step: 80, loss: 0.006777545902878046
step: 90, loss: 0.021291209384799004
step: 100, loss: 0.0033194830175489187
step: 110, loss: 0.030309803783893585
step: 120, loss: 0.005671079736202955
step: 130, loss: 0.014730509370565414
step: 140, loss: 0.05900612473487854
step: 150, loss: 0.13575954735279083
step: 160, loss: 0.004076753277331591
step: 170, loss: 0.002492345869541168
step: 180, loss: 0.04627276957035065
step: 190, loss: 0.0032229539938271046
step: 200, loss: 0.06455789506435394
step: 210, loss: 0.007765628397464752
step: 220, loss: 0.0011385519756004214
step: 230, loss: 0.002267594914883375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9663677130044843, f1=0.9659863945578231,
best_f1=0.9659863945578231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005944572854787111
step: 10, loss: 0.019489947706460953
step: 20, loss: 0.11377003788948059
step: 30, loss: 0.022419553250074387
step: 40, loss: 0.1412428468465805
step: 50, loss: 0.09143847972154617
step: 60, loss: 0.01166278962045908
step: 70, loss: 0.018140435218811035
step: 80, loss: 0.09009800106287003
step: 90, loss: 0.030996816232800484
step: 100, loss: 0.009145510382950306
step: 110, loss: 0.0020536629017442465
step: 120, loss: 0.0004891910357400775
step: 130, loss: 0.01102106086909771
step: 140, loss: 0.00811435654759407
step: 150, loss: 0.09036784619092941

```

```

step: 160, loss: 0.0011659699957817793
step: 170, loss: 0.0009711913880892098
step: 180, loss: 0.0019349217182025313
step: 190, loss: 0.05197662487626076
step: 200, loss: 0.0038411347195506096
step: 210, loss: 0.0005303464713506401
step: 220, loss: 0.0408325120806694
step: 230, loss: 0.041395604610443115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9695603156708005, f1=0.963718820861678,
best_f1=0.963718820861678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0062936157919466496
step: 10, loss: 0.0015581046463921666
step: 20, loss: 0.012334861792623997
step: 30, loss: 0.003332151100039482
step: 40, loss: 0.03086709976196289
step: 50, loss: 0.07639344781637192
step: 60, loss: 0.009801318868994713
step: 70, loss: 0.03219527751207352
step: 80, loss: 0.07271263003349304
step: 90, loss: 0.2998228073120117
step: 100, loss: 0.0737154632806778
step: 110, loss: 0.0445285402238369
step: 120, loss: 0.007349650375545025
step: 130, loss: 0.038085561245679855
step: 140, loss: 0.00298043224029243
step: 150, loss: 0.0005817132769152522
step: 160, loss: 0.0011754781007766724
step: 170, loss: 0.0008064270950853825
step: 180, loss: 0.15574167668819427
step: 190, loss: 0.0010193221969529986
step: 200, loss: 0.017676010727882385
step: 210, loss: 0.004365094937384129
step: 220, loss: 0.0017312464769929647
step: 230, loss: 0.031141983345150948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9798206278026906, f1=0.9763779527559054,

best_f1=0.9763779527559054

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

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

step: 0, loss: 0.001101407571695745

step: 10, loss: 0.0025441248435527086

step: 20, loss: 0.017556767910718918

step: 30, loss: 0.0007814753917045891

step: 40, loss: 0.0009931476088240743

step: 50, loss: 0.0005566374165937304

step: 60, loss: 0.014901780523359776

step: 70, loss: 0.0006525589851662517

step: 80, loss: 0.02065930888056755

step: 90, loss: 0.014253864996135235

step: 100, loss: 0.0003721113607753068

step: 110, loss: 0.000865817186422646

step: 120, loss: 0.006716479547321796

step: 130, loss: 0.0005375400651246309

step: 140, loss: 0.0004980633966624737

step: 150, loss: 0.031471651047468185

step: 160, loss: 0.0004751453234348446

step: 170, loss: 0.001354685053229332

step: 180, loss: 0.018052557483315468

step: 190, loss: 0.01555910799652338

step: 200, loss: 0.006851831916719675

step: 210, loss: 0.010891693644225597

step: 220, loss: 0.0009405399905517697

step: 230, loss: 0.002457326976582408

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

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

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

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=0.9819819819819819, f1=0.9774266365688488,

best_f1=0.9774266365688488

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

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

step: 0, loss: 0.00038997692172415555

step: 10, loss: 0.0032018469646573067

step: 20, loss: 0.0312671884894371

step: 30, loss: 0.004616730380803347

```

step: 40, loss: 0.0008067951421253383
step: 50, loss: 0.000888908514752984
step: 60, loss: 0.05133093520998955
step: 70, loss: 0.0024313514586538076
step: 80, loss: 0.0033238676842302084
step: 90, loss: 0.021197497844696045
step: 100, loss: 0.002645783359184861
step: 110, loss: 0.01236727461218834
step: 120, loss: 0.00043348243343643844
step: 130, loss: 0.0010382412001490593
step: 140, loss: 0.0005812265444546938
step: 150, loss: 0.00019336548575665802
step: 160, loss: 0.0005786109250038862
step: 170, loss: 0.002452895510941744
step: 180, loss: 0.001026140176691115
step: 190, loss: 0.028755521401762962
step: 200, loss: 0.0031790202483534813
step: 210, loss: 0.0006666811532340944
step: 220, loss: 0.03968261927366257
step: 230, loss: 0.0017197469715029001
/usr/local/lib/python3.7/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=0.9743016759776536, f1=0.976324689966178,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022228669840842485
step: 10, loss: 0.0015719181392341852
step: 20, loss: 0.024054184556007385
step: 30, loss: 0.00044802227057516575
step: 40, loss: 0.0011270256945863366
step: 50, loss: 0.04170899838209152
step: 60, loss: 0.0006220104405656457
step: 70, loss: 0.00047940018703229725
step: 80, loss: 0.0006742296973243356
step: 90, loss: 0.013160747475922108
step: 100, loss: 0.0023692050017416477
step: 110, loss: 0.0005165997426956892
step: 120, loss: 0.003928888123482466
step: 130, loss: 0.001287769409827888
step: 140, loss: 0.00031864861375652254
step: 150, loss: 0.0010655837832018733
step: 160, loss: 0.00025093439035117626
step: 170, loss: 0.0003618910850491375
step: 180, loss: 0.0003777925739996135
step: 190, loss: 0.016414297744631767

```

```

step: 200, loss: 0.0012692275922745466
step: 210, loss: 0.00619019428268075
step: 220, loss: 0.000557609018869698
step: 230, loss: 0.0022032298147678375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831271091113611, f1=0.9751131221719457,
best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003425563219934702
step: 10, loss: 0.003752436488866806
step: 20, loss: 0.002562862355262041
step: 30, loss: 0.15584678947925568
step: 40, loss: 0.0007369261002168059
step: 50, loss: 0.004797791596502066
step: 60, loss: 0.00041179475374519825
step: 70, loss: 0.00022587615239899606
step: 80, loss: 0.029797835275530815
step: 90, loss: 0.00042424985440447927
step: 100, loss: 0.0003983820497523993
step: 110, loss: 0.0013439644826576114
step: 120, loss: 0.001284634810872376
step: 130, loss: 0.0014311238192021847
step: 140, loss: 0.0001975902123376727
step: 150, loss: 0.0946756899356842
step: 160, loss: 0.0002062810235656798
step: 170, loss: 0.017263876274228096
step: 180, loss: 0.0004211720952298492
step: 190, loss: 0.00038233917439356446
step: 200, loss: 0.00030862880521453917
step: 210, loss: 0.0009671724983491004
step: 220, loss: 0.0003035583649761975
step: 230, loss: 0.0002072815696010366
/usr/local/lib/python3.7/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=0.9819413092550789, f1=0.9705215419501134,
best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0003258654032833874
step: 10, loss: 0.0002142571029253304
step: 20, loss: 0.0001809722016332671
step: 30, loss: 0.00019210166647098958
step: 40, loss: 0.0004316421691328287
step: 50, loss: 0.00029790267581120133
step: 60, loss: 0.0016113248420879245
step: 70, loss: 0.029304491356015205
step: 80, loss: 0.0001147252187365666
step: 90, loss: 0.028353942558169365
step: 100, loss: 0.00011527239257702604
step: 110, loss: 8.786458784015849e-05
step: 120, loss: 0.09601888805627823
step: 130, loss: 0.0004480089119169861
step: 140, loss: 0.00042021373519673944
step: 150, loss: 0.00019461064948700368
step: 160, loss: 0.00031416412093676627
step: 170, loss: 0.00040225297561846673
step: 180, loss: 0.00017542527348268777
step: 190, loss: 0.00013569796283263713
step: 200, loss: 0.0003857484261970967
step: 210, loss: 0.001463788328692317
step: 220, loss: 0.0021501751616597176
step: 230, loss: 0.0023808553814888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.983050847457627, f1=0.9668571428571429,
best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019362517923582345
step: 10, loss: 0.0004539660585578531
step: 20, loss: 0.00018711226584855467
step: 30, loss: 0.0001381569163640961
step: 40, loss: 0.0005510338814929128
step: 50, loss: 0.0002658805751707405
step: 60, loss: 0.0020472167525440454
step: 70, loss: 0.005430768709629774
step: 80, loss: 0.00014420801016967744
step: 90, loss: 0.0001630916667636484
step: 100, loss: 0.0007758966530673206
step: 110, loss: 0.0001463641383452341
step: 120, loss: 0.0002381520753260702
step: 130, loss: 0.0002627464709803462
step: 140, loss: 0.00013804453192278743
step: 150, loss: 0.00024817531812004745

```

```

step: 160, loss: 7.213024218799546e-05
step: 170, loss: 0.00011897517106262967
step: 180, loss: 0.015692852437496185
step: 190, loss: 9.458104614168406e-05
step: 200, loss: 0.0001875993621069938
step: 210, loss: 0.00018317479407414794
step: 220, loss: 0.0003012624802067876
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 6.917981954757124e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9841986455981941, f1=0.9751693002257337,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013007959933020175
step: 10, loss: 0.00016440004401374608
step: 20, loss: 0.019001029431819916
step: 30, loss: 0.00018711230950430036
step: 40, loss: 4.781875395565294e-05
step: 50, loss: 7.767775241518393e-05
step: 60, loss: 0.0019455060828477144
step: 70, loss: 0.00010809724335558712
step: 80, loss: 0.004369434434920549
step: 90, loss: 0.10514215379953384
step: 100, loss: 0.00024157235748134553
step: 110, loss: 0.0006327004521153867
step: 120, loss: 0.002213447354733944
step: 130, loss: 0.0002884744026232511
step: 140, loss: 0.0003490024828352034
step: 150, loss: 0.00028951591229997575
step: 160, loss: 0.0011505570728331804
step: 170, loss: 0.0019860791508108377
step: 180, loss: 0.0002716034068726003
step: 190, loss: 0.0001969961595023051
step: 200, loss: 0.0020564033184200525
step: 210, loss: 0.0004016267484985292
step: 220, loss: 0.0002935920492745936
step: 230, loss: 0.00020091434998903424
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9785310734463276,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020654867694247514
step: 10, loss: 0.00010803993791341782
step: 20, loss: 0.006362430285662413
step: 30, loss: 0.014650102704763412
step: 40, loss: 0.00011900862591573969
step: 50, loss: 0.0005863079568371177
step: 60, loss: 0.0001403539936291054
step: 70, loss: 0.00015407425235025585
step: 80, loss: 4.773922773892991e-05
step: 90, loss: 0.0005812564049847424
step: 100, loss: 6.312430195976049e-05
step: 110, loss: 5.7714558352017775e-05
step: 120, loss: 0.00020159628184046596
step: 130, loss: 6.765804573660716e-05
step: 140, loss: 0.00010281221329933032
step: 150, loss: 0.00017242533795069903
step: 160, loss: 0.001134852645918727
step: 170, loss: 0.00018381084373686463
step: 180, loss: 0.00016097720072139055
step: 190, loss: 0.002132116351276636
step: 200, loss: 7.721365545876324e-05
step: 210, loss: 0.046532873064279556
step: 220, loss: 0.01165207289159298
step: 230, loss: 0.0030373982153832912
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853438556933484, f1=0.9670079635949943,
best_f1=0.9670079635949943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006244083633646369
step: 10, loss: 0.0012901428854092956
step: 20, loss: 0.0015090641099959612
step: 30, loss: 4.118998549529351e-05
step: 40, loss: 9.157996100839227e-05
step: 50, loss: 0.010444235987961292
step: 60, loss: 8.990280912257731e-05

```

```

step: 70, loss: 0.0011580135906115174
step: 80, loss: 0.00016770666115917265
step: 90, loss: 5.835222327732481e-05
step: 100, loss: 9.983936615753919e-05
step: 110, loss: 0.0015320362290367484
step: 120, loss: 5.7722827477846295e-05
step: 130, loss: 0.00012722592509817332
step: 140, loss: 0.0023491729516535997
step: 150, loss: 7.950579311000183e-05
step: 160, loss: 0.004251136910170317
step: 170, loss: 8.434079063590616e-05
step: 180, loss: 0.009144796058535576
step: 190, loss: 6.888594361953437e-05
step: 200, loss: 0.0009880997240543365
step: 210, loss: 0.003737621707841754
step: 220, loss: 6.689973815809935e-05
step: 230, loss: 0.0014559603296220303
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9797297297297298,
best_f1=0.9670079635949943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.9437989218859e-05
step: 10, loss: 4.706201798398979e-05
step: 20, loss: 8.203287870856002e-05
step: 30, loss: 0.000908337882719934
step: 40, loss: 4.011930650449358e-05
step: 50, loss: 4.418577373144217e-05
step: 60, loss: 5.676390719600022e-05
step: 70, loss: 6.379646947607398e-05
step: 80, loss: 6.107131048338488e-05
step: 90, loss: 6.737122748745605e-05
step: 100, loss: 0.0006234762840904295
step: 110, loss: 9.444863098906353e-05
step: 120, loss: 3.747706796275452e-05
step: 130, loss: 0.008808416314423084
step: 140, loss: 0.00028129099518992007
step: 150, loss: 4.5005406718701124e-05
step: 160, loss: 0.00011674842244246975
step: 170, loss: 7.651471241842955e-05
step: 180, loss: 8.042873378144577e-05
step: 190, loss: 4.215242006466724e-05
step: 200, loss: 6.965991633478552e-05
step: 210, loss: 5.75650847167708e-05
step: 220, loss: 9.844297164818272e-05

```

```

step: 230, loss: 3.976103835157119e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9853438556933484, f1=0.9785794813979707,
best_f1=0.9670079635949943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006078869919292629
step: 10, loss: 6.222812226042151e-05
step: 20, loss: 6.41468504909426e-05
step: 30, loss: 5.084254007670097e-05
step: 40, loss: 3.582402132451534e-05
step: 50, loss: 8.229431114159524e-05
step: 60, loss: 0.0363374687731266
step: 70, loss: 8.933571371017024e-05
step: 80, loss: 0.0004509376303758472
step: 90, loss: 4.337722930358723e-05
step: 100, loss: 3.315727735753171e-05
step: 110, loss: 0.00025387544883415103
step: 120, loss: 0.018801374360919
step: 130, loss: 5.401501402957365e-05
step: 140, loss: 0.003636325476691127
step: 150, loss: 0.0001821759360609576
step: 160, loss: 0.00023202065494842827
step: 170, loss: 5.0021819333778694e-05
step: 180, loss: 6.238109199330211e-05
step: 190, loss: 0.0001544226543046534
step: 200, loss: 0.00018789473688229918
step: 210, loss: 0.013187567703425884
step: 220, loss: 5.777396654593758e-05
step: 230, loss: 0.00011909688328159973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9853438556933484, f1=0.9785794813979707,
best_f1=0.9670079635949943
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, 206.71it/s]
load_f1 = 0.9841986455981941
real_f1 = 0.9853438556933484
0it [00:00, ?it/s]/usr/local/lib/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, 190.87it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
 /usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
 An input tensor was not cuda.

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

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: 0.7218472361564636

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

step: 10, loss: 0.4450352191925049

step: 20, loss: 0.35046687722206116

step: 30, loss: 0.34971553087234497

step: 40, loss: 0.3592987060546875

step: 50, loss: 0.6634908318519592

step: 60, loss: 0.3801895081996918

step: 70, loss: 0.47288158535957336

step: 80, loss: 0.5044375061988831

step: 90, loss: 0.4449401795864105

step: 100, loss: 0.38798925280570984

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

step: 110, loss: 0.529499888420105

step: 120, loss: 0.6460742950439453

step: 130, loss: 0.5189985036849976

step: 140, loss: 0.30360502004623413

step: 150, loss: 0.17753984034061432

step: 160, loss: 0.3875259757041931

step: 170, loss: 0.33800581097602844

step: 180, loss: 0.12185690551996231

step: 190, loss: 0.37817293405532837

step: 200, loss: 0.15087640285491943

step: 210, loss: 0.089127317070961

step: 220, loss: 0.23954004049301147

step: 230, loss: 0.14445745944976807

step: 240, loss: 0.02089606411755085

step: 250, loss: 0.1837201863527298

step: 260, loss: 0.33117106556892395

step: 270, loss: 0.36834293603897095

step: 280, loss: 0.08194802701473236

step: 290, loss: 0.11035135388374329

step: 300, loss: 0.2948533594608307

step: 310, loss: 0.14755944907665253

step: 320, loss: 0.10895784944295883

step: 330, loss: 0.17909029126167297

step: 340, loss: 0.3554585576057434

step: 350, loss: 0.21016184985637665

step: 360, loss: 0.028276510536670685

step: 370, loss: 0.29547974467277527

step: 380, loss: 0.16834549605846405

step: 390, loss: 0.013907257467508316

step: 400, loss: 0.08118803799152374

step: 410, loss: 0.2682837247848511

```

step: 420, loss: 0.03443949297070503
step: 430, loss: 0.040374670177698135
step: 440, loss: 0.019060278311371803
step: 450, loss: 0.03812681511044502
step: 460, loss: 0.0686081051826477
step: 470, loss: 0.12934862077236176
step: 480, loss: 0.13229264318943024
step: 490, loss: 0.10037646442651749
step: 500, loss: 0.060632649809122086
step: 510, loss: 0.061346784234046936
step: 520, loss: 0.123466856777668
step: 530, loss: 0.16199086606502533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8978840846366145, f1=0.9043478260869564,
best_f1=0.9043478260869564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21321679651737213
step: 10, loss: 0.1895759403705597
step: 20, loss: 0.11109815537929535
step: 30, loss: 0.1155305877327919
step: 40, loss: 0.04934442788362503
step: 50, loss: 0.08960931748151779
step: 60, loss: 0.039591435343027115
step: 70, loss: 0.0929877758026123
step: 80, loss: 0.118870809674263
step: 90, loss: 0.05027610436081886
step: 100, loss: 0.25010454654693604
step: 110, loss: 0.019220925867557526
step: 120, loss: 0.1602724939584732
step: 130, loss: 0.011139927431941032
step: 140, loss: 0.03515784069895744
step: 150, loss: 0.015978680923581123
step: 160, loss: 0.07398976385593414
step: 170, loss: 0.10608028620481491
step: 180, loss: 0.08535732328891754
step: 190, loss: 0.024093052372336388
step: 200, loss: 0.2224835753440857
step: 210, loss: 0.0432840995490551
step: 220, loss: 0.002924420638009906
step: 230, loss: 0.18795397877693176

```

```

step: 240, loss: 0.019552335143089294
step: 250, loss: 0.0758906826376915
step: 260, loss: 0.14906272292137146
step: 270, loss: 0.09351623058319092
step: 280, loss: 0.08542609214782715
step: 290, loss: 0.030027693137526512
step: 300, loss: 0.05291682109236717
step: 310, loss: 0.04012652114033699
step: 320, loss: 0.0587298721075058
step: 330, loss: 0.1072741225361824
step: 340, loss: 0.1374187469482422
step: 350, loss: 0.03261996805667877
step: 360, loss: 0.06912286579608917
step: 370, loss: 0.037307046353816986
step: 380, loss: 0.20167022943496704
step: 390, loss: 0.013065174221992493
step: 400, loss: 0.05311190336942673
step: 410, loss: 0.015185672789812088
step: 420, loss: 0.03337307646870613
step: 430, loss: 0.06571197509765625
step: 440, loss: 0.021936872974038124
step: 450, loss: 0.045570291578769684
step: 460, loss: 0.08985497057437897
step: 470, loss: 0.08494098484516144
step: 480, loss: 0.015053086914122105
step: 490, loss: 0.043452877551317215
step: 500, loss: 0.007837077602744102
step: 510, loss: 0.01853228732943535
step: 520, loss: 0.38054120540618896
step: 530, loss: 0.21246933937072754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9380127620783956, f1=0.9353958143767062,
best_f1=0.9353958143767062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22533966600894928
step: 10, loss: 0.18924374878406525
step: 20, loss: 0.009610802866518497
step: 30, loss: 0.16361111402511597
step: 40, loss: 0.04485287144780159
step: 50, loss: 0.013692517764866352

```


step: 60, loss: 0.02254258096218109
step: 70, loss: 0.012766975909471512
step: 80, loss: 0.06385042518377304
step: 90, loss: 0.012588363140821457
step: 100, loss: 0.08466792106628418
step: 110, loss: 0.05284738540649414
step: 120, loss: 0.1763227880001068
step: 130, loss: 0.1439456045627594
step: 140, loss: 0.06076761707663536
step: 150, loss: 0.16538341343402863
step: 160, loss: 0.023522451519966125
step: 170, loss: 0.0031498305033892393
step: 180, loss: 0.010691576637327671
step: 190, loss: 0.06396318972110748
step: 200, loss: 0.026397928595542908
step: 210, loss: 0.028056243434548378
step: 220, loss: 0.1255718618631363
step: 230, loss: 0.04399368539452553
step: 240, loss: 0.0618944950401783
step: 250, loss: 0.05222934111952782
step: 260, loss: 0.07961004972457886
step: 270, loss: 0.01671976037323475
step: 280, loss: 0.06164293363690376
step: 290, loss: 0.018402189016342163
step: 300, loss: 0.10555419325828552
step: 310, loss: 0.1687842309474945
step: 320, loss: 0.029107509180903435
step: 330, loss: 0.02851765975356102
step: 340, loss: 0.016028618440032005
step: 350, loss: 0.1732412725687027
step: 360, loss: 0.013049261644482613
step: 370, loss: 0.22284992039203644
step: 380, loss: 0.024840395897626877
step: 390, loss: 0.04906155541539192
step: 400, loss: 0.28752073645591736
step: 410, loss: 0.03401947021484375
step: 420, loss: 0.02008466236293316
step: 430, loss: 0.02478565275669098
step: 440, loss: 0.22462862730026245
step: 450, loss: 0.06482064723968506
step: 460, loss: 0.0708325058221817
step: 470, loss: 0.018317468464374542
step: 480, loss: 0.18052925169467926
step: 490, loss: 0.05405895784497261
step: 500, loss: 0.03934788703918457
step: 510, loss: 0.05992501974105835
step: 520, loss: 0.012054193764925003
step: 530, loss: 0.027349211275577545

```
/usr/local/lib/python3.7/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=0.9355586462679647, f1=0.9376163873370577,
best_f1=0.9353958143767062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041211385279893875
step: 10, loss: 0.0027914727106690407
step: 20, loss: 0.0926392674446106
step: 30, loss: 0.10901828110218048
step: 40, loss: 0.011739034205675125
step: 50, loss: 0.0346413291990757
step: 60, loss: 0.03172510117292404
step: 70, loss: 0.07973989099264145
step: 80, loss: 0.12576711177825928
step: 90, loss: 0.032332003116607666
step: 100, loss: 0.028557278215885162
step: 110, loss: 0.15369293093681335
step: 120, loss: 0.009521348401904106
step: 130, loss: 0.2713776230812073
step: 140, loss: 0.09396902471780777
step: 150, loss: 0.010474671609699726
step: 160, loss: 0.015482783317565918
step: 170, loss: 0.03667484596371651
step: 180, loss: 0.10842911154031754
step: 190, loss: 0.06553877890110016
step: 200, loss: 0.16921278834342957
step: 210, loss: 0.027784831821918488
step: 220, loss: 0.059782978147268295
step: 230, loss: 0.03330700099468231
step: 240, loss: 0.011431381106376648
step: 250, loss: 0.01745789498090744
step: 260, loss: 0.03093717433512211
step: 270, loss: 0.1665160357952118
step: 280, loss: 0.005164232105016708
step: 290, loss: 0.014735802076756954
step: 300, loss: 0.008603904396295547
step: 310, loss: 0.005340615753084421
step: 320, loss: 0.17653034627437592
step: 330, loss: 0.011334333568811417
step: 340, loss: 0.004497089888900518
step: 350, loss: 0.06351379305124283
step: 360, loss: 0.014665907248854637
step: 370, loss: 0.0029742680490016937
step: 380, loss: 0.0027633614372462034
step: 390, loss: 0.0027568882796913385
```

```

step: 400, loss: 0.04977897182106972
step: 410, loss: 0.004596485756337643
step: 420, loss: 0.016908369958400726
step: 430, loss: 0.06737644225358963
step: 440, loss: 0.0032892681192606688
step: 450, loss: 0.06854911148548126
step: 460, loss: 0.06230747699737549
step: 470, loss: 0.004954187665134668
step: 480, loss: 0.06670641899108887
step: 490, loss: 0.2591116726398468
step: 500, loss: 0.15801119804382324
step: 510, loss: 0.19626466929912567
step: 520, loss: 0.018808340653777122
step: 530, loss: 0.11478250473737717
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9431345353675451, f1=0.9360331339162449,
best_f1=0.9360331339162449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004211848601698875
step: 10, loss: 0.03741637244820595
step: 20, loss: 0.01096362341195345
step: 30, loss: 0.1321595460176468
step: 40, loss: 0.00810330081731081
step: 50, loss: 0.10365847498178482
step: 60, loss: 0.06413906812667847
step: 70, loss: 0.022310009226202965
step: 80, loss: 0.005105442367494106
step: 90, loss: 0.054665323346853256
step: 100, loss: 0.12227831780910492
step: 110, loss: 0.007292021531611681
step: 120, loss: 0.17689883708953857
step: 130, loss: 0.017751174047589302
step: 140, loss: 0.10357896983623505
step: 150, loss: 0.02896999381482601
step: 160, loss: 0.17232109606266022
step: 170, loss: 0.017467405647039413
step: 180, loss: 0.027912456542253494
step: 190, loss: 0.006193654611706734
step: 200, loss: 0.005938936024904251
step: 210, loss: 0.0005740422639064491

```

```

step: 220, loss: 0.018326932564377785
step: 230, loss: 0.013641119003295898
step: 240, loss: 0.0190159622579813
step: 250, loss: 0.09529096633195877
step: 260, loss: 0.010367950424551964
step: 270, loss: 0.009471853263676167
step: 280, loss: 0.011560799553990364
step: 290, loss: 0.0495145246386528
step: 300, loss: 0.0827193558216095
step: 310, loss: 0.007471853401511908
step: 320, loss: 0.2046431452035904
step: 330, loss: 0.08580543100833893
step: 340, loss: 0.012942817993462086
step: 350, loss: 0.0013488373951986432
step: 360, loss: 0.015890710055828094
step: 370, loss: 0.004537200555205345
step: 380, loss: 0.0021563339978456497
step: 390, loss: 0.03300242871046066
step: 400, loss: 0.027119016274809837
step: 410, loss: 0.10889545828104019
step: 420, loss: 0.20015966892242432
step: 430, loss: 0.08656821399927139
step: 440, loss: 0.002483286429196596
step: 450, loss: 0.03671219199895859
step: 460, loss: 0.04166923835873604
step: 470, loss: 0.046911995857954025
step: 480, loss: 0.013411476276814938
step: 490, loss: 0.06888244301080704
step: 500, loss: 0.09574724733829498
step: 510, loss: 0.009801882319152355
step: 520, loss: 0.10106595605611801
step: 530, loss: 0.01987258903682232
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9446768944676894, f1=0.9370629370629371,
best_f1=0.9370629370629371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0711178183555603
step: 10, loss: 0.001416526036337018
step: 20, loss: 0.0059336284175515175
step: 30, loss: 0.0138310007750988

```

step: 40, loss: 0.0064102052710950375
step: 50, loss: 0.011486247181892395
step: 60, loss: 0.04207605496048927
step: 70, loss: 0.005187225993722677
step: 80, loss: 0.0036526822950690985
step: 90, loss: 0.020664682611823082
step: 100, loss: 0.008597356267273426
step: 110, loss: 0.04590006545186043
step: 120, loss: 0.17026346921920776
step: 130, loss: 0.01897701434791088
step: 140, loss: 0.016654152423143387
step: 150, loss: 0.004253614693880081
step: 160, loss: 0.09113603085279465
step: 170, loss: 0.007047978229820728
step: 180, loss: 0.007267246022820473
step: 190, loss: 0.02618582919239998
step: 200, loss: 0.009640556760132313
step: 210, loss: 0.00345069682225585
step: 220, loss: 0.0019219864625483751
step: 230, loss: 0.006515792105346918
step: 240, loss: 0.013772383332252502
step: 250, loss: 0.06676594913005829
step: 260, loss: 0.004305845592170954
step: 270, loss: 0.014263136312365532
step: 280, loss: 0.0053257266990840435
step: 290, loss: 0.002842992078512907
step: 300, loss: 0.040584806352853775
step: 310, loss: 0.11776751279830933
step: 320, loss: 0.0031749638728797436
step: 330, loss: 0.005745674483478069
step: 340, loss: 0.006712361238896847
step: 350, loss: 0.002348392503336072
step: 360, loss: 0.030893614515662193
step: 370, loss: 0.032244257628917694
step: 380, loss: 0.00034951427369378507
step: 390, loss: 0.05986018851399422
step: 400, loss: 0.01938544027507305
step: 410, loss: 0.004215757828205824
step: 420, loss: 0.007052230183035135
step: 430, loss: 0.10204050689935684
step: 440, loss: 0.003458896651864052
step: 450, loss: 0.16470025479793549
step: 460, loss: 0.0012990923132747412
step: 470, loss: 0.0049508302472531796
step: 480, loss: 0.009420079179108143
step: 490, loss: 0.02342267334461212
step: 500, loss: 0.011445865035057068
step: 510, loss: 0.11582779139280319

```

step: 520, loss: 0.0013370042433962226
step: 530, loss: 0.005462448112666607
/usr/local/lib/python3.7/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=0.9408450704225352, f1=0.9320113314447592,
best_f1=0.9370629370629371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004272126592695713
step: 10, loss: 0.001955309882760048
step: 20, loss: 0.010567850433290005
step: 30, loss: 0.03114449977874756
step: 40, loss: 0.004750151187181473
step: 50, loss: 0.025192011147737503
step: 60, loss: 0.003758006729185581
step: 70, loss: 0.006463390775024891
step: 80, loss: 0.0011055328650400043
step: 90, loss: 0.0013775682309642434
step: 100, loss: 0.011460493318736553
step: 110, loss: 0.0011664677876979113
step: 120, loss: 0.0020703021436929703
step: 130, loss: 0.0035368395037949085
step: 140, loss: 0.013416019268333912
step: 150, loss: 0.0022750282660126686
step: 160, loss: 0.00036158118746243417
step: 170, loss: 0.0033449542243033648
step: 180, loss: 0.05641481280326843
step: 190, loss: 0.01697680726647377
step: 200, loss: 0.0004313515673857182
step: 210, loss: 0.004012168850749731
step: 220, loss: 0.00215522525832057
step: 230, loss: 0.0004236277600284666
step: 240, loss: 0.02718108519911766
step: 250, loss: 0.06683211028575897
step: 260, loss: 0.00445997528731823
step: 270, loss: 0.00256976205855608
step: 280, loss: 0.009976888075470924
step: 290, loss: 0.01004627626389265
step: 300, loss: 0.00230921758338809
step: 310, loss: 0.0019505281234160066
step: 320, loss: 0.020821867510676384
step: 330, loss: 0.016759764403104782
step: 340, loss: 0.0518844872713089
step: 350, loss: 0.004379157908260822
step: 360, loss: 0.13927355408668518
step: 370, loss: 0.0018993985140696168

```

```

step: 380, loss: 0.014555907808244228
step: 390, loss: 0.005307428538799286
step: 400, loss: 0.01301006879657507
step: 410, loss: 0.0012185070663690567
step: 420, loss: 0.04117341339588165
step: 430, loss: 0.0011283355997875333
step: 440, loss: 0.01939447410404682
step: 450, loss: 0.004678661003708839
step: 460, loss: 0.0014351304853335023
step: 470, loss: 0.14570142328739166
step: 480, loss: 0.00964651070535183
step: 490, loss: 0.009749931283295155
step: 500, loss: 0.002109163673594594
step: 510, loss: 0.004294717684388161
step: 520, loss: 0.03151369467377663
step: 530, loss: 0.0015060527948662639
/usr/local/lib/python3.7/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=0.9367088607594937, f1=0.933774834437086,
best_f1=0.9370629370629371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09102477133274078
step: 10, loss: 0.001844184473156929
step: 20, loss: 0.0027455235831439495
step: 30, loss: 0.007167752366513014
step: 40, loss: 0.00217923684976995
step: 50, loss: 0.000849299191031605
step: 60, loss: 0.00200764206238091
step: 70, loss: 0.013904472813010216
step: 80, loss: 0.0010000813053920865
step: 90, loss: 0.006331645883619785
step: 100, loss: 0.0023084981366991997
step: 110, loss: 0.0022798292338848114
step: 120, loss: 0.0035377938766032457
step: 130, loss: 0.1116957888007164
step: 140, loss: 0.006088673137128353
step: 150, loss: 0.0007352252141572535
step: 160, loss: 0.008280963636934757
step: 170, loss: 0.029935088008642197
step: 180, loss: 0.07237833738327026
step: 190, loss: 0.053141895681619644
step: 200, loss: 0.02841828018426895
step: 210, loss: 0.13079535961151123
step: 220, loss: 0.003179488005116582
step: 230, loss: 0.07506720721721649

```

```

step: 240, loss: 0.015696773305535316
step: 250, loss: 0.00047388908569701016
step: 260, loss: 0.00014251480752136558
step: 270, loss: 0.0018583218334242702
step: 280, loss: 0.0005807241541333497
step: 290, loss: 0.005402066744863987
step: 300, loss: 0.003973288927227259
step: 310, loss: 0.0068602911196649075
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 320, loss: 0.016404101625084877
step: 330, loss: 0.0005249694222584367
step: 340, loss: 0.006809300277382135
step: 350, loss: 0.0028481115587055683
step: 360, loss: 0.00530332513153553
step: 370, loss: 0.17864276468753815
step: 380, loss: 0.00022637877555098385
step: 390, loss: 0.007709058467298746
step: 400, loss: 0.004263447597622871
step: 410, loss: 0.019165143370628357
step: 420, loss: 0.009421068243682384
step: 430, loss: 0.006444780621677637
step: 440, loss: 0.010978877544403076
step: 450, loss: 0.0005620736628770828
step: 460, loss: 0.018727106973528862
step: 470, loss: 0.08340169489383698
step: 480, loss: 0.0014714369317516685
step: 490, loss: 0.1293870061635971
step: 500, loss: 0.0008503345889039338
step: 510, loss: 0.006478935480117798
step: 520, loss: 0.002938374411314726
step: 530, loss: 0.0037133251316845417
/usr/local/lib/python3.7/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=0.9421028253821214, f1=0.9348127600554785,
best_f1=0.9370629370629371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001366654527373612
step: 10, loss: 0.008165627717971802
step: 20, loss: 0.0004973990726284683
step: 30, loss: 0.03275945037603378
step: 40, loss: 0.0004351807583589107
step: 50, loss: 0.0005295445444062352
step: 60, loss: 0.0037551256828010082
step: 70, loss: 0.0005874300259165466
step: 80, loss: 0.027709951624274254

```



```
step: 90, loss: 0.05391532555222511
step: 100, loss: 0.007735266350209713
step: 110, loss: 0.04837945103645325
step: 120, loss: 0.018061354756355286
step: 130, loss: 0.014129262417554855
step: 140, loss: 0.0034118068870157003
step: 150, loss: 0.0016185055719688535
step: 160, loss: 0.0008724104845896363
step: 170, loss: 0.002402795944362879
step: 180, loss: 0.003037268528714776
step: 190, loss: 0.0004630143230315298
step: 200, loss: 0.0034593490418046713
step: 210, loss: 0.0013525018002837896
step: 220, loss: 0.004633234813809395
step: 230, loss: 0.0005803584936074913
step: 240, loss: 0.0012840612325817347
step: 250, loss: 0.026501979678869247
step: 260, loss: 0.001814045011997223
step: 270, loss: 0.006743403617292643
step: 280, loss: 0.0042462474666535854
step: 290, loss: 0.004811248276382685
step: 300, loss: 0.0003860622819047421
step: 310, loss: 0.1818545013666153
step: 320, loss: 0.00022398184228222817
step: 330, loss: 0.0007782069151289761
step: 340, loss: 0.03456305339932442
step: 350, loss: 0.05535350367426872
step: 360, loss: 0.017229942604899406
step: 370, loss: 0.0018639853224158287
step: 380, loss: 0.0007818673620931804
step: 390, loss: 0.00025250299950130284
step: 400, loss: 0.0037172159645706415
step: 410, loss: 0.0006316349608823657
step: 420, loss: 0.0003806512977462262
step: 430, loss: 0.011072791181504726
step: 440, loss: 0.12579473853111267
step: 450, loss: 0.03748786449432373
step: 460, loss: 0.0006500571616925299
step: 470, loss: 0.0003372998326085508
step: 480, loss: 0.0008062291308306158
step: 490, loss: 0.0666549876332283
step: 500, loss: 0.0009824112057685852
step: 510, loss: 0.010916613973677158
step: 520, loss: 0.003127642907202244
step: 530, loss: 0.07130677998065948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9452830188679245, f1=0.9407372841810545,
best_f1=0.9407372841810545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000584048277232796
step: 10, loss: 0.0030184437055140734
step: 20, loss: 0.00018234435992781073
step: 30, loss: 0.0319828987121582
step: 40, loss: 0.00016835196583997458
step: 50, loss: 0.0035512084141373634
step: 60, loss: 0.0031147818081080914
step: 70, loss: 0.008928600698709488
step: 80, loss: 0.0003384602605365217
step: 90, loss: 0.0028632511384785175
step: 100, loss: 0.0007928419508971274
step: 110, loss: 0.05414503812789917
step: 120, loss: 0.0012641667854040861
step: 130, loss: 0.020235203206539154
step: 140, loss: 0.011515199206769466
step: 150, loss: 0.003206107532605529
step: 160, loss: 0.07709358632564545
step: 170, loss: 0.0026755628641694784
step: 180, loss: 0.008638626895844936
step: 190, loss: 0.00049652683082968
step: 200, loss: 0.0020671524107456207
step: 210, loss: 0.0014291024999693036
step: 220, loss: 0.017100665718317032
step: 230, loss: 0.003860642435029149
step: 240, loss: 0.0011708203237503767
step: 250, loss: 0.010691745206713676
step: 260, loss: 0.0460846871137619
step: 270, loss: 0.02235439606010914
step: 280, loss: 0.018084494397044182
step: 290, loss: 0.0008917954401113093
step: 300, loss: 0.028736334294080734
step: 310, loss: 0.0049724397249519825
step: 320, loss: 0.01067261677235365
step: 330, loss: 0.005010609515011311
step: 340, loss: 0.011978239752352238
step: 350, loss: 0.00032980297692120075
step: 360, loss: 0.00029307964723557234
step: 370, loss: 0.0016284097218886018
step: 380, loss: 0.0010696660028770566

```

```

step: 390, loss: 0.0019475955050438643
step: 400, loss: 0.014328058809041977
step: 410, loss: 0.00019091530703008175
step: 420, loss: 0.07234882563352585
step: 430, loss: 0.0004222100833430886
step: 440, loss: 0.00022132520098239183
step: 450, loss: 0.019177226349711418
step: 460, loss: 0.00035988425952382386
step: 470, loss: 0.0027534635737538338
step: 480, loss: 0.0013603271218016744
step: 490, loss: 0.008112771436572075
step: 500, loss: 0.001892912550829351
step: 510, loss: 9.88682295428589e-05
step: 520, loss: 0.03800303861498833
step: 530, loss: 0.00829513743519783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9463276836158193, f1=0.9404536862003782,
best_f1=0.9404536862003782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004995265626348555
step: 10, loss: 0.0003237247292418033
step: 20, loss: 0.0022375821135938168
step: 30, loss: 0.0001659314875723794
step: 40, loss: 0.00015148251259233803
step: 50, loss: 0.009176960214972496
step: 60, loss: 0.001793975243344903
step: 70, loss: 0.016185754910111427
step: 80, loss: 0.0016242857091128826
step: 90, loss: 0.0063398065976798534
step: 100, loss: 0.00364328152500093
step: 110, loss: 0.0002747733087744564
step: 120, loss: 0.02616041898727417
step: 130, loss: 0.011200721375644207
step: 140, loss: 0.012824979610741138
step: 150, loss: 0.004815001040697098
step: 160, loss: 0.0012920969165861607
step: 170, loss: 0.003416384570300579
step: 180, loss: 0.00743162352591753
step: 190, loss: 0.00159907853230834
step: 200, loss: 0.0015340606914833188

```

```

step: 210, loss: 0.0003085736825596541
step: 220, loss: 0.048721011728048325
step: 230, loss: 0.0011689650127664208
step: 240, loss: 0.005285849794745445
step: 250, loss: 0.07654904574155807
step: 260, loss: 0.008371797390282154
step: 270, loss: 0.0013543193927034736
step: 280, loss: 0.008513690903782845
step: 290, loss: 0.004275166429579258
step: 300, loss: 0.0011628068750724196
step: 310, loss: 0.1256195455789566
step: 320, loss: 0.024376288056373596
step: 330, loss: 6.390824273694307e-05
step: 340, loss: 0.006077633239328861
step: 350, loss: 0.000725268735550344
step: 360, loss: 0.00461539626121521
step: 370, loss: 0.0019435319118201733
step: 380, loss: 0.0027033421210944653
step: 390, loss: 0.0027381866239011288
step: 400, loss: 0.0002987308253068477
step: 410, loss: 0.000249205797445029
step: 420, loss: 0.0006518674199469388
step: 430, loss: 0.0018421766581013799
step: 440, loss: 0.00030733412131667137
step: 450, loss: 0.0022396626882255077
step: 460, loss: 0.004736908245831728
step: 470, loss: 0.005654909648001194
step: 480, loss: 0.001074911910109222
step: 490, loss: 0.00013795788981951773
step: 500, loss: 0.0004159384116064757
step: 510, loss: 0.000524340255651623
step: 520, loss: 0.002439414616674185
step: 530, loss: 0.045026201754808426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9464944649446495, f1=0.9442653155228006,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00718949967995286
step: 10, loss: 0.001050750259310007
step: 20, loss: 0.004262845031917095

```

step: 30, loss: 0.00013517668412532657
step: 40, loss: 0.0008943771244958043
step: 50, loss: 9.744987619342282e-05
step: 60, loss: 0.011190869845449924
step: 70, loss: 0.012776925228536129
step: 80, loss: 0.06112239509820938
step: 90, loss: 0.08260250091552734
step: 100, loss: 0.07592429965734482
step: 110, loss: 0.08396878838539124
step: 120, loss: 0.001776816789060831
step: 130, loss: 0.0015022066654637456
step: 140, loss: 0.0016741219442337751
step: 150, loss: 0.0001375297870254144
step: 160, loss: 0.0075167324393987656
step: 170, loss: 0.0005266225198283792
step: 180, loss: 0.0002482171112205833
step: 190, loss: 0.000711682834662497
step: 200, loss: 0.0004564027476590127
step: 210, loss: 0.0006341783446259797
step: 220, loss: 0.00015498962602578104
step: 230, loss: 0.016305863857269287
step: 240, loss: 0.00014169764472171664
step: 250, loss: 4.390817412058823e-05
step: 260, loss: 0.035684388130903244
step: 270, loss: 0.008660203777253628
step: 280, loss: 0.0002922863350249827
step: 290, loss: 0.0046190680004656315
step: 300, loss: 0.002247751923277974
step: 310, loss: 0.0013192427577450871
step: 320, loss: 0.0002905359142459929
step: 330, loss: 0.02776039019227028
step: 340, loss: 0.021172992885112762
step: 350, loss: 0.006614520214498043
step: 360, loss: 0.0004416831652633846
step: 370, loss: 0.0006824107258580625
step: 380, loss: 0.00020547310123220086
step: 390, loss: 0.024357903748750687
step: 400, loss: 0.008880197070538998
step: 410, loss: 0.03859001025557518
step: 420, loss: 5.461479304358363e-05
step: 430, loss: 0.00038546687574125826
step: 440, loss: 0.018570642918348312
step: 450, loss: 0.029801087453961372
step: 460, loss: 0.008446426130831242
step: 470, loss: 0.002535605803132057
step: 480, loss: 0.01024688221514225
step: 490, loss: 0.0009710690937936306
step: 500, loss: 0.0008214993285946548

```

step: 510, loss: 0.00047934349277056754
step: 520, loss: 9.148885146714747e-05
step: 530, loss: 0.0002427553408779204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9481132075471699, f1=0.9431123648330982,
best_f1=0.9431123648330982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005121954716742039
step: 10, loss: 0.0005025489954277873
step: 20, loss: 0.00014422889216803014
step: 30, loss: 0.0010192099725827575
step: 40, loss: 0.002281761495396495
step: 50, loss: 0.004048701375722885
step: 60, loss: 0.001144580659456551
step: 70, loss: 0.005654430482536554
step: 80, loss: 0.002555516781285405
step: 90, loss: 0.0003177466569468379
step: 100, loss: 0.0015848701586946845
step: 110, loss: 0.010760932229459286
step: 120, loss: 0.015408664010465145
step: 130, loss: 0.0002943467115983367
step: 140, loss: 0.0002448578889016062
step: 150, loss: 0.002297826111316681
step: 160, loss: 0.01050649955868721
step: 170, loss: 0.014273744076490402
step: 180, loss: 0.0010108109563589096
step: 190, loss: 0.0009111523977480829
step: 200, loss: 0.0002901724656112492
step: 210, loss: 0.00015593213902320713
step: 220, loss: 0.005099721718579531
step: 230, loss: 0.0029748629312962294
step: 240, loss: 0.0020260822493582964
step: 250, loss: 0.00896613858640194
step: 260, loss: 0.001503176405094564
step: 270, loss: 0.02456524781882763
step: 280, loss: 0.00013207730080466717
step: 290, loss: 0.00011397357593523338
step: 300, loss: 6.939517334103584e-05
step: 310, loss: 0.00038273606332950294
step: 320, loss: 6.474000838352367e-05

```

```

step: 330, loss: 0.009071226231753826
step: 340, loss: 0.03534883260726929
step: 350, loss: 0.00013673539797309786
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 360, loss: 0.09959807991981506
step: 370, loss: 0.0797789916396141
step: 380, loss: 0.0014925134601071477
step: 390, loss: 0.0001963943795999512
step: 400, loss: 0.00135913270059973
step: 410, loss: 0.0005498317186720669
step: 420, loss: 0.001449831761419773
step: 430, loss: 0.011826764792203903
step: 440, loss: 8.537193207303062e-05
step: 450, loss: 0.0007171383476816118
step: 460, loss: 0.0013059631455689669
step: 470, loss: 0.0028309389017522335
step: 480, loss: 9.544930071569979e-05
step: 490, loss: 0.000940938713029027
step: 500, loss: 0.001518629607744515
step: 510, loss: 0.00035554825444705784
step: 520, loss: 0.00029575766529887915
step: 530, loss: 0.0013531639706343412
/usr/local/lib/python3.7/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=0.946927374301676, f1=0.9416160672582905,
best_f1=0.9431123648330982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018151644326280802
step: 10, loss: 0.0015362352132797241
step: 20, loss: 0.0001703704328974709
step: 30, loss: 0.0007351632812060416
step: 40, loss: 0.0004972314927726984
step: 50, loss: 0.00010851123079191893
step: 60, loss: 0.005471146199852228
step: 70, loss: 0.004130973480641842
step: 80, loss: 0.00045862956903874874
step: 90, loss: 5.343741941032931e-05
step: 100, loss: 9.680179209681228e-05
step: 110, loss: 0.0006629868876188993
step: 120, loss: 0.00014541689597535878
step: 130, loss: 5.159743159310892e-05
step: 140, loss: 0.03128480166196823
step: 150, loss: 0.0009931606473401189
step: 160, loss: 0.009804398752748966
step: 170, loss: 0.010200198739767075

```

```

step: 180, loss: 0.00018344268028158695
step: 190, loss: 0.00020924682030454278
step: 200, loss: 3.291057873866521e-05
step: 210, loss: 0.044557806104421616
step: 220, loss: 0.0004459169285837561
step: 230, loss: 0.00020421094086486846
step: 240, loss: 0.006076454650610685
step: 250, loss: 0.003771076677367091
step: 260, loss: 0.00013678921095561236
step: 270, loss: 0.00046311141340993345
step: 280, loss: 0.00020715613209176809
step: 290, loss: 0.000420448457589373
step: 300, loss: 4.439177064341493e-05
step: 310, loss: 0.0005140174180269241
step: 320, loss: 0.00019433956185821444
step: 330, loss: 0.0009356018854305148
step: 340, loss: 0.000259222841123119
step: 350, loss: 8.587542833993211e-05
step: 360, loss: 0.0005754121229983866
step: 370, loss: 0.006503336597234011
step: 380, loss: 0.0029257526621222496
step: 390, loss: 0.0007522693485952914
step: 400, loss: 0.002447971608489752
step: 410, loss: 0.00036320119397714734
step: 420, loss: 8.964516746345907e-05
step: 430, loss: 0.002709187800064683
step: 440, loss: 0.00043986935634166
step: 450, loss: 0.0004448778345249593
step: 460, loss: 0.054169490933418274
step: 470, loss: 0.00019838263688143343
step: 480, loss: 0.00013249259791336954
step: 490, loss: 0.0017904448322951794
step: 500, loss: 0.005215305369347334
step: 510, loss: 0.01629256270825863
step: 520, loss: 0.00018899438146036118
step: 530, loss: 0.0004258947737980634
/usr/local/lib/python3.7/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=0.9446254071661238, f1=0.9416705552963137,
best_f1=0.9431123648330982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013408481609076262
step: 10, loss: 0.0011041434481739998
step: 20, loss: 0.00030219898326322436
step: 30, loss: 0.00717130396515131

```


step: 40, loss: 4.732962406706065e-05
step: 50, loss: 0.0008643860928714275
step: 60, loss: 0.010167034342885017
step: 70, loss: 0.00017790998390410095
step: 80, loss: 0.00012976520520169288
step: 90, loss: 0.0004891256103292108
step: 100, loss: 7.272006769198924e-05
step: 110, loss: 0.0006095041171647608
step: 120, loss: 4.788943624589592e-05
step: 130, loss: 0.00014031703176442534
step: 140, loss: 5.0504084356362e-05
step: 150, loss: 0.0009723567054606974
step: 160, loss: 4.064754830324091e-05
step: 170, loss: 0.0022427092771977186
step: 180, loss: 0.020754382014274597
step: 190, loss: 0.01325273048132658
step: 200, loss: 0.0041008684784173965
step: 210, loss: 0.002112712012603879
step: 220, loss: 0.00040136129246093333
step: 230, loss: 0.0062713781371712685
step: 240, loss: 0.0012326579308137298
step: 250, loss: 4.5241387852001935e-05
step: 260, loss: 0.00021149791427887976
step: 270, loss: 0.00012846132449340075
step: 280, loss: 0.0008081034175120294
step: 290, loss: 0.0003093225823249668
step: 300, loss: 5.712880374630913e-05
step: 310, loss: 0.05407257005572319
step: 320, loss: 5.991376383462921e-05
step: 330, loss: 0.01064604613929987
step: 340, loss: 0.00013306421169545501
step: 350, loss: 0.006829036865383387
step: 360, loss: 0.0020246515050530434
step: 370, loss: 0.0018705726834014058
step: 380, loss: 0.0005106347962282598
step: 390, loss: 9.636043978389353e-05
step: 400, loss: 0.0011884488631039858
step: 410, loss: 0.0019823850598186255
step: 420, loss: 0.00010132947500096634
step: 430, loss: 6.539861351484433e-05
step: 440, loss: 0.0003668325662147254
step: 450, loss: 0.0002881918044295162
step: 460, loss: 0.003294538240879774
step: 470, loss: 2.4187735107261688e-05
step: 480, loss: 0.0013329049106687307
step: 490, loss: 0.0179392509162426
step: 500, loss: 0.04371214285492897
step: 510, loss: 0.00017064505664166063

```

step: 520, loss: 5.624878758681007e-05
step: 530, loss: 7.318461575778201e-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=0.9448082319925164, f1=0.9418386491557222,
best_f1=0.9431123648330982
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, 254.62it/s]
load_f1 = 0.9456221198156683
real_f1 = 0.9461077844311377
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 207.08it/s]

```

2 BASELINE

2.1 BASELINE STRUCTURED

2.1.1 Beer - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

```

After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4658133089542389
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=0.2777777777777778, f1=0.25, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4107532501220703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.35294117647058826, f1=0.36734693877551017,
best_f1=0.36734693877551017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3942429721355438
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4126984126984127, f1=0.40816326530612246,
best_f1=0.40816326530612246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26882028579711914
/usr/local/lib/python3.7/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=0.345679012345679, f1=0.40625, best_f1=0.40816326530612246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.272389680147171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.41935483870967744, f1=0.36, best_f1=0.36
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22063083946704865
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.5106382978723404, f1=0.4, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.4590015411376953
/usr/local/lib/python3.7/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=0.2947368421052632, f1=0.3023255813953489, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4919452965259552
/usr/local/lib/python3.7/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=0.37209302325581395, f1=0.3414634146341463, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3163464665412903
/usr/local/lib/python3.7/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=0.40625, f1=0.37037037037037035, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37101006507873535
/usr/local/lib/python3.7/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=0.4905660377358491, f1=0.42553191489361697, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3762083053588867
/usr/local/lib/python3.7/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=0.5090909090909091, f1=0.4545454545454545, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2309122234582901
/usr/local/lib/python3.7/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=0.45901639344262296, f1=0.40816326530612246, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.2857477366924286
/usr/local/lib/python3.7/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=0.5, f1=0.5, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22271963953971863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5116279069767441, 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: 0.23926717042922974
/usr/local/lib/python3.7/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=0.5116279069767441, f1=0.5, best_f1=0.5
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, 125635.83it/s]
/usr/local/lib/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 = 0.5000000000000001
real_f1 = 0.5116279069767441
0it [00:00, ?it/s]/usr/local/lib/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, 175.72it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

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

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

```

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

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

step: 0, loss: 0.587480902671814

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

step: 10, loss: 0.5039095282554626

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

step: 20, loss: 0.4698675870895386

step: 30, loss: 0.3380136787891388

step: 40, loss: 0.3347395956516266

step: 50, loss: 0.6067061424255371

step: 60, loss: 0.4926571846008301

step: 70, loss: 0.4281964898109436

step: 80, loss: 0.576214075088501

step: 90, loss: 0.48822730779647827

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

step: 100, loss: 0.4971965551376343

step: 110, loss: 0.5878348350524902

step: 120, loss: 0.42297208309173584

step: 130, loss: 0.35983484983444214

step: 140, loss: 0.39288994669914246

step: 150, loss: 0.5937826633453369

step: 160, loss: 0.4823971390724182

step: 170, loss: 0.49438613653182983

step: 180, loss: 0.38569438457489014

step: 190, loss: 0.23128609359264374

```

step: 200, loss: 0.2621867060661316
step: 210, loss: 0.22391928732395172
step: 220, loss: 0.24803043901920319
step: 230, loss: 0.048627421259880066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.939655172413793, f1=0.9397590361445783,
best_f1=0.9397590361445783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1045503243803978
step: 10, loss: 0.10331667214632034
step: 20, loss: 0.017292067408561707
step: 30, loss: 0.06249409168958664
step: 40, loss: 0.018456384539604187
step: 50, loss: 0.007364640943706036
step: 60, loss: 0.014799879863858223
step: 70, loss: 0.050809815526008606
step: 80, loss: 0.024728119373321533
step: 90, loss: 0.02877052314579487
step: 100, loss: 0.155246764421463
step: 110, loss: 0.27480658888816833
step: 120, loss: 0.09436548501253128
step: 130, loss: 0.01211319025605917
step: 140, loss: 0.004947267007082701
step: 150, loss: 0.023791423067450523
step: 160, loss: 0.07626287639141083
step: 170, loss: 0.006993809714913368
step: 180, loss: 0.03398338705301285
step: 190, loss: 0.008434628136456013
step: 200, loss: 0.016689080744981766
step: 210, loss: 0.002766626188531518
step: 220, loss: 0.0014402390224859118
step: 230, loss: 0.006826768163591623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9784335981838819, f1=0.977116704805492,

```

```

best_f1=0.977116704805492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020356841385364532
step: 10, loss: 0.006731148809194565
step: 20, loss: 0.0012775760842487216
step: 30, loss: 0.01661866530776024
step: 40, loss: 0.04009043797850609
step: 50, loss: 0.10078399628400803
step: 60, loss: 0.004659728612750769
step: 70, loss: 0.03945391625165939
step: 80, loss: 0.00276023056358099
step: 90, loss: 0.009703382849693298
step: 100, loss: 0.009773329831659794
step: 110, loss: 0.11530651152133942
step: 120, loss: 0.0031842906028032303
step: 130, loss: 0.005708147771656513
step: 140, loss: 0.0008584039751440287
step: 150, loss: 0.024232076480984688
step: 160, loss: 0.018749777227640152
step: 170, loss: 0.02124658040702343
step: 180, loss: 0.017709704115986824
step: 190, loss: 0.02639688365161419
step: 200, loss: 0.02734210714697838
step: 210, loss: 0.0023654985707253218
step: 220, loss: 0.011886381544172764
step: 230, loss: 0.038113072514534
/usr/local/lib/python3.7/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=0.976324689966178, f1=0.9714937286202965,
best_f1=0.977116704805492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05776609107851982
step: 10, loss: 0.01328064315021038
step: 20, loss: 0.0010550949955359101
step: 30, loss: 0.0007259762496687472
step: 40, loss: 0.14277763664722443
step: 50, loss: 0.007672763429582119
step: 60, loss: 0.012111591175198555
step: 70, loss: 0.01802140660583973
step: 80, loss: 0.0021908211056143045
step: 90, loss: 0.022330621257424355
step: 100, loss: 0.0056680249981582165
step: 110, loss: 0.0007429267861880362

```

```

step: 120, loss: 0.004940832499414682
step: 130, loss: 0.0010841485345736146
step: 140, loss: 0.0007380704628303647
step: 150, loss: 0.08205081522464752
step: 160, loss: 0.03135973960161209
step: 170, loss: 0.08811440318822861
step: 180, loss: 0.128753662109375
step: 190, loss: 0.003080332186073065
step: 200, loss: 0.0028215653728693724
step: 210, loss: 0.0038658450357615948
step: 220, loss: 0.00048645969945937395
step: 230, loss: 0.0031992767471820116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853438556933484, f1=0.9806157354618015,
best_f1=0.9806157354618015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014004685916006565
step: 10, loss: 0.0017998102121055126
step: 20, loss: 0.0016352171078324318
step: 30, loss: 0.0012238684576004744
step: 40, loss: 0.030091172084212303
step: 50, loss: 0.0005144437891431153
step: 60, loss: 0.003750080009922385
step: 70, loss: 0.0006456365226767957
step: 80, loss: 0.186090350151062
step: 90, loss: 0.024332556873559952
step: 100, loss: 0.00048788345884531736
step: 110, loss: 0.007301920093595982
step: 120, loss: 0.0026692282408475876
step: 130, loss: 0.009724346920847893
step: 140, loss: 0.017890485003590584
step: 150, loss: 0.00913957692682743
step: 160, loss: 0.028944138437509537
step: 170, loss: 0.0007206156733445823
step: 180, loss: 0.0012345408322289586
step: 190, loss: 0.1015024334192276
step: 200, loss: 0.00273827719502151
step: 210, loss: 0.0018192180432379246
step: 220, loss: 0.0006611442659050226
step: 230, loss: 0.0029199623968452215

```

```

/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9841269841269841,
best_f1=0.9806157354618015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032446099794469774
step: 10, loss: 0.0021264057140797377
step: 20, loss: 0.00072811683639884
step: 30, loss: 0.0005462815170176327
step: 40, loss: 0.00022471141710411757
step: 50, loss: 0.0003313279594294727
step: 60, loss: 0.04897218570113182
step: 70, loss: 0.00037109272670932114
step: 80, loss: 0.000729667954146862
step: 90, loss: 0.00448905723169446
step: 100, loss: 0.00783711951225996
step: 110, loss: 0.017571985721588135
step: 120, loss: 0.0006355282966978848
step: 130, loss: 0.03873680159449577
step: 140, loss: 0.004876908380538225
step: 150, loss: 0.00030072644585743546
step: 160, loss: 0.0015366876032203436
step: 170, loss: 0.0009433146915398538
step: 180, loss: 0.006261139176785946
step: 190, loss: 0.00026707423967309296
step: 200, loss: 0.0012478200951591134
step: 210, loss: 0.012367953546345234
step: 220, loss: 0.004154715687036514
step: 230, loss: 0.002101401099935174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9808342728297633, f1=0.9806598407281,
best_f1=0.9806157354618015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018839426338672638
step: 10, loss: 0.0009302771068178117
step: 20, loss: 0.0004004507209174335
step: 30, loss: 0.0003140574845019728
step: 40, loss: 0.0007882344070822
step: 50, loss: 0.002333143725991249
step: 60, loss: 0.00019729952327907085
step: 70, loss: 0.00043727681622840464

```

```

step: 80, loss: 0.01761048473417759
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.05789422243833542
step: 100, loss: 0.002618660219013691
step: 110, loss: 0.00029986441950313747
step: 120, loss: 0.0007157897343859076
step: 130, loss: 0.005368965212255716
step: 140, loss: 0.0002666419604793191
step: 150, loss: 0.015176406130194664
step: 160, loss: 0.0028051454573869705
step: 170, loss: 0.15452460944652557
step: 180, loss: 0.016439232975244522
step: 190, loss: 0.00029570574406534433
step: 200, loss: 0.006057858932763338
step: 210, loss: 0.0033872390631586313
step: 220, loss: 0.0009263165993615985
step: 230, loss: 0.0018097175052389503
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.9785794813979707,
best_f1=0.9806157354618015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020614927634596825
step: 10, loss: 0.00230276002548635
step: 20, loss: 0.0010627054143697023
step: 30, loss: 0.0014781325589865446
step: 40, loss: 0.0029476811178028584
step: 50, loss: 0.0007524591637775302
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 60, loss: 0.0007983465911820531
step: 70, loss: 0.00024703252711333334
step: 80, loss: 0.019748510792851448
step: 90, loss: 0.0006018223357386887
step: 100, loss: 0.0013049474218860269
step: 110, loss: 0.0219242125749588
step: 120, loss: 0.0004106522537767887
step: 130, loss: 0.006700304336845875
step: 140, loss: 0.001859612762928009
step: 150, loss: 0.012765470892190933
step: 160, loss: 0.008134093135595322
step: 170, loss: 0.12490952759981155
step: 180, loss: 0.00025124946841970086
step: 190, loss: 0.0008855875348672271
step: 200, loss: 0.0003263467806391418
step: 210, loss: 0.0020268578082323074

```

```

step: 220, loss: 0.004497931804507971
step: 230, loss: 0.002573443576693535
/usr/local/lib/python3.7/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=0.9785794813979707, f1=0.9783845278725825,
best_f1=0.9806157354618015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005691631231456995
step: 10, loss: 0.0008687472436577082
step: 20, loss: 0.0023940529208630323
step: 30, loss: 0.0012601586058735847
step: 40, loss: 0.00046593352453783154
step: 50, loss: 0.0014619212597608566
step: 60, loss: 0.0004039327905047685
step: 70, loss: 0.04475530609488487
step: 80, loss: 0.0012737340293824673
step: 90, loss: 0.017643267288804054
step: 100, loss: 0.000453352986369282
step: 110, loss: 0.0007692009676247835
step: 120, loss: 0.010309425182640553
step: 130, loss: 0.00042932317592203617
step: 140, loss: 0.0003035622357856482
step: 150, loss: 0.01847342774271965
step: 160, loss: 0.003117393236607313
step: 170, loss: 0.0006248405552469194
step: 180, loss: 0.00023336586309596896
step: 190, loss: 0.0001323810574831441
step: 200, loss: 0.008917412720620632
step: 210, loss: 0.13326546549797058
step: 220, loss: 0.0005568611668422818
step: 230, loss: 0.000252499186899513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9819413092550789, f1=0.9796380090497738,
best_f1=0.9806157354618015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000654106552246958
step: 10, loss: 0.000515245134010911
step: 20, loss: 0.00033880688715726137
step: 30, loss: 0.00015993916895240545
step: 40, loss: 0.0004280635912436992
step: 50, loss: 0.00023360010527539998

```

```

step: 60, loss: 0.0012250388972461224
step: 70, loss: 0.012072873301804066
step: 80, loss: 0.000212438142625615
step: 90, loss: 0.00019228064047638327
step: 100, loss: 0.00016019021859392524
step: 110, loss: 0.00014905039279256016
step: 120, loss: 8.978414552984759e-05
step: 130, loss: 0.00016112213779706508
step: 140, loss: 0.000251161924097687
step: 150, loss: 0.001417935243807733
step: 160, loss: 7.842384366085753e-05
step: 170, loss: 0.0001417166495230049
step: 180, loss: 0.0012038436252623796
step: 190, loss: 8.968933980213478e-05
step: 200, loss: 0.00012208471889607608
step: 210, loss: 0.007168652955442667
step: 220, loss: 0.01752244494855404
step: 230, loss: 0.0005707976524718106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9875706214689265, f1=0.9818594104308391,
best_f1=0.9818594104308391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.159161759773269e-05
step: 10, loss: 0.00023225245240610093
step: 20, loss: 0.00012810515181627125
step: 30, loss: 0.0001240992423845455
step: 40, loss: 5.604911711998284e-05
step: 50, loss: 0.0013963504461571574
step: 60, loss: 0.003851430956274271
step: 70, loss: 8.69364885147661e-05
step: 80, loss: 0.0001314301771344617
step: 90, loss: 0.2159041464328766
step: 100, loss: 0.00015957506548147649
step: 110, loss: 0.0006654191529378295
step: 120, loss: 0.00019501241331454366
step: 130, loss: 0.00021483855380211025
step: 140, loss: 0.002353711985051632
step: 150, loss: 0.0003231608134228736
step: 160, loss: 0.00032567125163041055
step: 170, loss: 0.00023247624631039798

```



```

step: 180, loss: 0.0001818769087549299
step: 190, loss: 0.00013954759924672544
step: 200, loss: 0.0003324735153000802
step: 210, loss: 0.00016129568393807858
step: 220, loss: 0.03150907903909683
step: 230, loss: 0.00042165053309872746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.983050847457627,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010752028319984674
step: 10, loss: 0.00024261753424070776
step: 20, loss: 0.0006225364049896598
step: 30, loss: 0.009688228368759155
step: 40, loss: 0.00024379015667364
step: 50, loss: 0.00041663029696792364
step: 60, loss: 0.00022407484357245266
step: 70, loss: 0.00021366108558140695
step: 80, loss: 0.000108943204395473
step: 90, loss: 0.00025022297631949186
step: 100, loss: 0.00010057692270493135
step: 110, loss: 0.00013483039219863713
step: 120, loss: 0.00021270943398121744
step: 130, loss: 0.0003217354533262551
step: 140, loss: 0.0006447919877246022
step: 150, loss: 0.00037908877129666507
step: 160, loss: 0.0003997072053607553
step: 170, loss: 0.00028885481879115105
step: 180, loss: 0.00023308108211494982
step: 190, loss: 0.0009949536761268973
step: 200, loss: 0.0002839747758116573
step: 210, loss: 0.0002713132998906076
step: 220, loss: 0.0023819440975785255
step: 230, loss: 0.0008164949249476194
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9819413092550789,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000304721383145079
step: 10, loss: 0.0001784775813575834
step: 20, loss: 0.00033537851413711905
step: 30, loss: 8.031471952563152e-05
step: 40, loss: 0.00025824145996011794
step: 50, loss: 0.0005499265971593559
step: 60, loss: 0.00020402958034537733
step: 70, loss: 0.00012160180631326512
step: 80, loss: 0.0001443003275198862
step: 90, loss: 0.00019685820734594017
step: 100, loss: 0.0001961856905836612
step: 110, loss: 0.002355600008741021
step: 120, loss: 0.0003150725387968123
step: 130, loss: 0.00042861519614234567
step: 140, loss: 0.00015659509517718107
step: 150, loss: 7.287097105290741e-05
step: 160, loss: 0.0010036922758445144
step: 170, loss: 0.00017959393153432757
step: 180, loss: 0.0011450998717918992
step: 190, loss: 0.00012516634888015687
step: 200, loss: 6.359146937029436e-05
step: 210, loss: 0.0002756344329100102
step: 220, loss: 0.0003363537252880633
step: 230, loss: 0.0003233129100408405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9865168539325843, f1=0.9819413092550789,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008827458950690925
step: 10, loss: 0.00011731481208698824
step: 20, loss: 0.00021968904184177518
step: 30, loss: 0.0002767622936517
step: 40, loss: 8.934027573559433e-05
step: 50, loss: 0.00013715609384234995
step: 60, loss: 0.0003312825574539602
step: 70, loss: 0.0003034012916032225
step: 80, loss: 0.00010872309212572873
step: 90, loss: 0.0003312390181235969
step: 100, loss: 0.00025851474492810667
step: 110, loss: 0.000125858437968418
step: 120, loss: 4.9363232392352074e-05
step: 130, loss: 0.0005176890990696847

```

```

step: 140, loss: 0.0002798849600367248
step: 150, loss: 0.00012275842891540378
step: 160, loss: 0.0001315735571552068
step: 170, loss: 9.910509106703103e-05
step: 180, loss: 0.0002197493304265663
step: 190, loss: 0.00012433150550350547
step: 200, loss: 0.00033932796213775873
step: 210, loss: 7.330612424993888e-05
step: 220, loss: 9.856160613708198e-05
step: 230, loss: 8.805451216176152e-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=0.9854423292273236, f1=0.9831271091113611,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007477807812392712
step: 10, loss: 0.00013083001249469817
step: 20, loss: 0.00017814144666772336
step: 30, loss: 0.0001868167018983513
step: 40, loss: 4.906698814011179e-05
step: 50, loss: 9.314011549577117e-05
step: 60, loss: 0.026975097134709358
step: 70, loss: 0.00016322676674462855
step: 80, loss: 0.0003893919347319752
step: 90, loss: 9.849356138147414e-05
step: 100, loss: 8.165701001416892e-05
step: 110, loss: 0.001242452533915639
step: 120, loss: 0.0006014711107127368
step: 130, loss: 8.551998325856403e-05
step: 140, loss: 0.00040264124982059
step: 150, loss: 0.0002152534871129319
step: 160, loss: 0.0008233049884438515
step: 170, loss: 4.467175676836632e-05
step: 180, loss: 0.0017873261822387576
step: 190, loss: 9.395920642418787e-05
step: 200, loss: 0.00041317768045701087
step: 210, loss: 0.00017797204782254994
step: 220, loss: 8.415993943344802e-05
step: 230, loss: 0.00011317828466417268
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9831271091113611,
best_f1=0.983050847457627
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 [00:13, 182.14it/s]

load_f1 = 0.9898534385569334

real_f1 = 0.9876265466816648

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

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

4398it [00:26, 169.11it/s]

2.1.3 DBLP-GoogleScholar - Running the matcher

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

```

--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model

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

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 448kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 2.83MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 1.81MB/s]
Downloading: 100% 501M/501M [00:14<00:00, 34.1MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly

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

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6199216246604919
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41932249069213867
step: 20, loss: 0.341363787651062
step: 30, loss: 0.3572133183479309
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.3538145422935486
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 50, loss: 1.1408671140670776
step: 60, loss: 0.3507251441478729
step: 70, loss: 0.29065877199172974
step: 80, loss: 0.27225059270858765
step: 90, loss: 0.26164934039115906
step: 100, loss: 0.2951529026031494
step: 110, loss: 0.11753363907337189
step: 120, loss: 0.1476854383945465
step: 130, loss: 0.4772719740867615
step: 140, loss: 0.12106967717409134
step: 150, loss: 0.0708668902516365
step: 160, loss: 0.2133818119764328
step: 170, loss: 0.21965014934539795
step: 180, loss: 0.08545301109552383
step: 190, loss: 0.03933696821331978
step: 200, loss: 0.0567902997136116
step: 210, loss: 0.06369183212518692
step: 220, loss: 0.06412239372730255
step: 230, loss: 0.19890853762626648
step: 240, loss: 0.13010656833648682
step: 250, loss: 0.05662127584218979
step: 260, loss: 0.23137791454792023
step: 270, loss: 0.21477800607681274
step: 280, loss: 0.0816168561577797
step: 290, loss: 0.0694238543510437
step: 300, loss: 0.11671297997236252
step: 310, loss: 0.251627653837204
step: 320, loss: 0.11528608947992325
step: 330, loss: 0.05822193622589111
step: 340, loss: 0.4109538793563843
step: 350, loss: 0.11618859320878983
step: 360, loss: 0.06964355707168579
step: 370, loss: 0.05211824178695679
step: 380, loss: 0.1279812753200531
```

```

step: 390, loss: 0.023571960628032684
step: 400, loss: 0.0740874707698822
step: 410, loss: 0.3181842863559723
step: 420, loss: 0.055621348321437836
step: 430, loss: 0.052846238017082214
step: 440, loss: 0.02283412590622902
step: 450, loss: 0.013149005360901356
step: 460, loss: 0.028425250202417374
step: 470, loss: 0.0757148489356041
step: 480, loss: 0.20533901453018188
step: 490, loss: 0.22613465785980225
step: 500, loss: 0.07615827769041061
step: 510, loss: 0.08809755742549896
step: 520, loss: 0.04455764219164848
step: 530, loss: 0.20581400394439697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9314553990610329, f1=0.9259431765253843,
best_f1=0.9259431765253843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05120714381337166
step: 10, loss: 0.0639725923538208
step: 20, loss: 0.047990582883358
step: 30, loss: 0.07715733349323273
step: 40, loss: 0.024012351408600807
step: 50, loss: 0.09031949937343597
step: 60, loss: 0.035161279141902924
step: 70, loss: 0.03638112172484398
step: 80, loss: 0.03304094076156616
step: 90, loss: 0.021791256964206696
step: 100, loss: 0.19126005470752716
step: 110, loss: 0.02383950911462307
step: 120, loss: 0.09468347579240799
step: 130, loss: 0.024062883108854294
step: 140, loss: 0.015702305361628532
step: 150, loss: 0.016284054145216942
step: 160, loss: 0.05128677934408188
step: 170, loss: 0.09224069863557816
step: 180, loss: 0.0880563035607338
step: 190, loss: 0.03131767734885216
step: 200, loss: 0.2088528275489807

```

```

step: 210, loss: 0.06374191492795944
step: 220, loss: 0.003129425458610058
step: 230, loss: 0.045665886253118515
step: 240, loss: 0.053160108625888824
step: 250, loss: 0.07647936791181564
step: 260, loss: 0.06599961221218109
step: 270, loss: 0.06959562003612518
step: 280, loss: 0.08702843636274338
step: 290, loss: 0.08230633288621902
step: 300, loss: 0.04153527691960335
step: 310, loss: 0.05453421548008919
step: 320, loss: 0.12767978012561798
step: 330, loss: 0.14977861940860748
step: 340, loss: 0.043806128203868866
step: 350, loss: 0.00207309378311038
step: 360, loss: 0.1075638011097908
step: 370, loss: 0.07160152494907379
step: 380, loss: 0.16830165684223175
step: 390, loss: 0.011076647788286209
step: 400, loss: 0.1479795128107071
step: 410, loss: 0.056058403104543686
step: 420, loss: 0.05102372169494629
step: 430, loss: 0.4918368458747864
step: 440, loss: 0.11237446218729019
step: 450, loss: 0.07671769708395004
step: 460, loss: 0.020295750349760056
step: 470, loss: 0.13312752544879913
step: 480, loss: 0.04806320369243622
step: 490, loss: 0.0858171358704567
step: 500, loss: 0.01805396005511284
step: 510, loss: 0.03662753850221634
step: 520, loss: 0.3518858551979065
step: 530, loss: 0.03703857958316803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9460227272727273, f1=0.9352720450281427,
best_f1=0.9352720450281427
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07219067215919495
step: 10, loss: 0.05580689013004303
step: 20, loss: 0.00859611015766859

```


step: 30, loss: 0.07862623035907745
step: 40, loss: 0.051409538835287094
step: 50, loss: 0.009014246053993702
step: 60, loss: 0.03901268169283867
step: 70, loss: 0.008238743059337139
step: 80, loss: 0.10960201174020767
step: 90, loss: 0.006074290722608566
step: 100, loss: 0.043404340744018555
step: 110, loss: 0.09134730696678162
step: 120, loss: 0.13414934277534485
step: 130, loss: 0.10662179440259933
step: 140, loss: 0.0076587325893342495
step: 150, loss: 0.01762758567929268
step: 160, loss: 0.02618785947561264
step: 170, loss: 0.004495114553719759
step: 180, loss: 0.03310808166861534
step: 190, loss: 0.008719896897673607
step: 200, loss: 0.05490870773792267
step: 210, loss: 0.05230340361595154
step: 220, loss: 0.11239909380674362
step: 230, loss: 0.0328594408929348
step: 240, loss: 0.058760400861501694
step: 250, loss: 0.12229858338832855
step: 260, loss: 0.11106835305690765
step: 270, loss: 0.001566237537190318
step: 280, loss: 0.03793441504240036
step: 290, loss: 0.012257146649062634
step: 300, loss: 0.2110820710659027
step: 310, loss: 0.06438245624303818
step: 320, loss: 0.031248804181814194
step: 330, loss: 0.027126846835017204
step: 340, loss: 0.0026392778381705284
step: 350, loss: 0.020723821595311165
step: 360, loss: 0.018415037542581558
step: 370, loss: 0.05892253667116165
step: 380, loss: 0.07320517301559448
step: 390, loss: 0.025700315833091736
step: 400, loss: 0.10116752237081528
step: 410, loss: 0.12035204470157623
step: 420, loss: 0.005873224698007107
step: 430, loss: 0.012709859758615494
step: 440, loss: 0.24069000780582428
step: 450, loss: 0.020886551588773727
step: 460, loss: 0.1652570366859436
step: 470, loss: 0.003450753865763545
step: 480, loss: 0.07656793296337128
step: 490, loss: 0.012683792039752007
step: 500, loss: 0.02141781710088253

```

step: 510, loss: 0.0540204793214798
step: 520, loss: 0.0021745169069617987
step: 530, loss: 0.09945836663246155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.951356407857811, f1=0.9433085501858736,
best_f1=0.9433085501858736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016490191221237183
step: 10, loss: 0.004805610049515963
step: 20, loss: 0.1475774347782135
step: 30, loss: 0.03025186061859131
step: 40, loss: 0.007010405883193016
step: 50, loss: 0.05016029253602028
step: 60, loss: 0.03499988466501236
step: 70, loss: 0.034717779606580734
step: 80, loss: 0.12579090893268585
step: 90, loss: 0.019447239115834236
step: 100, loss: 0.009655529633164406
step: 110, loss: 0.1468873918056488
step: 120, loss: 0.0025272336788475513
step: 130, loss: 0.012850113213062286
step: 140, loss: 0.011228809133172035
step: 150, loss: 0.006938375066965818
step: 160, loss: 0.007638717070221901
step: 170, loss: 0.018174072727560997
step: 180, loss: 0.08355754613876343
step: 190, loss: 0.041883453726768494
step: 200, loss: 0.008530726656317711
step: 210, loss: 0.0032304839696735144
step: 220, loss: 0.005620485637336969
step: 230, loss: 0.0052840495482087135
step: 240, loss: 0.023063842207193375
step: 250, loss: 0.06266793608665466
step: 260, loss: 0.009805076755583286
step: 270, loss: 0.1106441542506218
step: 280, loss: 0.0031545062083750963
step: 290, loss: 0.05551823973655701
step: 300, loss: 0.00623589102178812
step: 310, loss: 0.001735601108521223
step: 320, loss: 0.1364550143480301

```

```

step: 330, loss: 0.02400074526667595
step: 340, loss: 0.002070787362754345
step: 350, loss: 0.03953557461500168
step: 360, loss: 0.01133421715348959
step: 370, loss: 0.013726689852774143
step: 380, loss: 0.0047961934469640255
step: 390, loss: 0.011526994407176971
step: 400, loss: 0.013074221089482307
step: 410, loss: 0.008882773108780384
step: 420, loss: 0.002969768363982439
step: 430, loss: 0.003856262192130089
step: 440, loss: 0.001508902758359909
step: 450, loss: 0.01244447287172079
step: 460, loss: 0.014239798299968243
step: 470, loss: 0.003581662429496646
step: 480, loss: 0.004873559810221195
step: 490, loss: 0.002394565613940358
step: 500, loss: 0.022493386641144753
step: 510, loss: 0.11687858402729034
step: 520, loss: 0.013488576747477055
step: 530, loss: 0.12445726245641708
/usr/local/lib/python3.7/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=0.950668510834486, f1=0.9409070087036188,
best_f1=0.9433085501858736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040396228432655334
step: 10, loss: 0.006849465426057577
step: 20, loss: 0.0017639667494222522
step: 30, loss: 0.12863008677959442
step: 40, loss: 0.005027074366807938
step: 50, loss: 0.006494046188890934
step: 60, loss: 0.009044049307703972
step: 70, loss: 0.08514897525310516
step: 80, loss: 0.008567877113819122
step: 90, loss: 0.06212225556373596
step: 100, loss: 0.048427291214466095
step: 110, loss: 0.03502945229411125
step: 120, loss: 0.2164374589920044
step: 130, loss: 0.0406842902302742
step: 140, loss: 0.019816908985376358
step: 150, loss: 0.005114202853292227
step: 160, loss: 0.003097479697316885
step: 170, loss: 0.025499334558844566
step: 180, loss: 0.0046690115705132484

```

```

step: 190, loss: 0.01427287794649601
step: 200, loss: 0.007747814524918795
step: 210, loss: 0.0015705975238233805
step: 220, loss: 0.003749832510948181
step: 230, loss: 0.0040190075524151325
step: 240, loss: 0.009106685407459736
step: 250, loss: 0.11616644263267517
step: 260, loss: 0.015583176165819168
step: 270, loss: 0.016221532598137856
step: 280, loss: 0.030547501519322395
step: 290, loss: 0.0019309383351355791
step: 300, loss: 0.014883778989315033
step: 310, loss: 0.06629234552383423
step: 320, loss: 0.12871569395065308
step: 330, loss: 0.018948964774608612
step: 340, loss: 0.01565689407289028
step: 350, loss: 0.001280717784538865
step: 360, loss: 0.0001947209530044347
step: 370, loss: 0.005397281143814325
step: 380, loss: 0.0025799162685871124
step: 390, loss: 0.003756769699975848
step: 400, loss: 0.013875038363039494
step: 410, loss: 0.12525033950805664
step: 420, loss: 0.0863075777888298
step: 430, loss: 0.07035236805677414
step: 440, loss: 0.015126275829970837
step: 450, loss: 0.0049228304997086525
step: 460, loss: 0.03829242289066315
step: 470, loss: 0.03577914461493492
step: 480, loss: 0.023500358685851097
step: 490, loss: 0.016465172171592712
step: 500, loss: 0.08322221785783768
step: 510, loss: 0.0017360096098855138
step: 520, loss: 0.11090613156557083
step: 530, loss: 0.0005256989970803261
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9514563106796116, f1=0.9388505747126438,
best_f1=0.9388505747126438
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017170461360365152

```

step: 10, loss: 0.0006417492404580116
step: 20, loss: 0.0006457063718698919
step: 30, loss: 0.007374868728220463
step: 40, loss: 0.0012059956789016724
step: 50, loss: 0.06760042160749435
step: 60, loss: 0.13217762112617493
step: 70, loss: 0.07459966093301773
step: 80, loss: 0.0009505494381301105
step: 90, loss: 0.025405386462807655
step: 100, loss: 0.07198462635278702
step: 110, loss: 0.0034711535554379225
step: 120, loss: 0.10680367052555084
step: 130, loss: 0.0004916171892546117
step: 140, loss: 0.0024397445376962423
step: 150, loss: 0.0023107966408133507
step: 160, loss: 0.012116427533328533
step: 170, loss: 0.005981765687465668
step: 180, loss: 0.005521886516362429
step: 190, loss: 0.03434503823518753
step: 200, loss: 0.07272624224424362
step: 210, loss: 0.013906941749155521
step: 220, loss: 0.00272663333453238
step: 230, loss: 0.009318974800407887
step: 240, loss: 0.020477386191487312
step: 250, loss: 0.0030403488781303167
step: 260, loss: 0.001439437735825777
step: 270, loss: 0.0008021796238608658
step: 280, loss: 0.005901539232581854
step: 290, loss: 0.0025368602946400642
step: 300, loss: 0.0008181497105397284
step: 310, loss: 0.06982001662254333
step: 320, loss: 0.00015857072139624506
step: 330, loss: 0.012574221007525921
step: 340, loss: 0.0005717705935239792
step: 350, loss: 0.0024482952430844307
step: 360, loss: 0.0014199698343873024
step: 370, loss: 0.03922770172357559
step: 380, loss: 0.0008824141114018857
step: 390, loss: 0.00021930047660134733
step: 400, loss: 0.06297945976257324
step: 410, loss: 0.006749724969267845
step: 420, loss: 0.010917022824287415
step: 430, loss: 0.010860170237720013
step: 440, loss: 0.00128056644462049
step: 450, loss: 0.12440352141857147
step: 460, loss: 0.008790507912635803
step: 470, loss: 0.00885525718331337
step: 480, loss: 0.005894084926694632

```

step: 490, loss: 0.05969458818435669
step: 500, loss: 0.003181144595146179
step: 510, loss: 0.22855710983276367
step: 520, loss: 0.00633582565933466
step: 530, loss: 0.05325581133365631
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9540768509840675, f1=0.937471051412691,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004728012252599001
step: 10, loss: 0.0021502648014575243
step: 20, loss: 0.0027863874565809965
step: 30, loss: 0.034396324306726456
step: 40, loss: 0.0038876303005963564
step: 50, loss: 0.006735881790518761
step: 60, loss: 0.004806156270205975
step: 70, loss: 0.001725492300465703
step: 80, loss: 0.0035857255570590496
step: 90, loss: 3.6903435102431104e-05
step: 100, loss: 8.702767081558704e-05
step: 110, loss: 0.00019100411736872047
step: 120, loss: 0.07682207971811295
step: 130, loss: 0.003654958214610815
step: 140, loss: 0.0004984194529242814
step: 150, loss: 0.08998569846153259
step: 160, loss: 0.001906254910863936
step: 170, loss: 0.0030515764374285936
step: 180, loss: 0.007832486182451248
step: 190, loss: 0.0024130723904818296
step: 200, loss: 0.000877337355632335
step: 210, loss: 0.0006960812024772167
step: 220, loss: 0.00046439727884717286
step: 230, loss: 0.00012665647955145687
step: 240, loss: 0.0005823078099638224
step: 250, loss: 0.038789670914411545
step: 260, loss: 0.09938729554414749
step: 270, loss: 0.001068967510946095
step: 280, loss: 0.01306965947151184
step: 290, loss: 0.014618685469031334
step: 300, loss: 0.0027543199248611927

```

```

step: 310, loss: 0.0014079411048442125
step: 320, loss: 0.02857639268040657
step: 330, loss: 0.0025568243581801653
step: 340, loss: 0.0020217078272253275
step: 350, loss: 0.0006100113969296217
step: 360, loss: 0.003142581321299076
step: 370, loss: 0.02029629983007908
step: 380, loss: 0.0019054441945627332
step: 390, loss: 0.000421366305090487
step: 400, loss: 0.00776659743860364
step: 410, loss: 0.018218262121081352
step: 420, loss: 0.012034132145345211
step: 430, loss: 0.0008478460367769003
step: 440, loss: 0.0008706182125024498
step: 450, loss: 0.025433210656046867
step: 460, loss: 0.00015725503908470273
step: 470, loss: 0.016377070918679237
step: 480, loss: 0.0010766605846583843
step: 490, loss: 0.004322986118495464
step: 500, loss: 0.0006384849548339844
step: 510, loss: 0.00031142687657848
step: 520, loss: 0.0008230968960560858
step: 530, loss: 0.0006577749154530466
/usr/local/lib/python3.7/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=0.9535864978902954, f1=0.9438515081206497,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019800167065113783
step: 10, loss: 0.000759742280934006
step: 20, loss: 0.07669010013341904
step: 30, loss: 0.0002035547950072214
step: 40, loss: 0.00020164487068541348
step: 50, loss: 0.008496449328958988
step: 60, loss: 0.0023620796855539083
step: 70, loss: 0.002773156389594078
step: 80, loss: 0.016230585053563118
step: 90, loss: 0.0004713061498478055
step: 100, loss: 0.025220943614840508
step: 110, loss: 0.0007764040492475033
step: 120, loss: 0.0017122238641604781
step: 130, loss: 0.002396051771938801
step: 140, loss: 0.00716800382360816
step: 150, loss: 0.00027286840486340225
step: 160, loss: 0.00012722867541015148

```

```

step: 170, loss: 0.0017712527187541127
step: 180, loss: 0.00015791808255016804
step: 190, loss: 0.00021404445578809828
step: 200, loss: 0.0024239409249275923
step: 210, loss: 0.050539106130599976
step: 220, loss: 0.003429330186918378
step: 230, loss: 0.05330801010131836
step: 240, loss: 0.0006130895344540477
step: 250, loss: 0.0009224169189110398
step: 260, loss: 0.00020340975606814027
step: 270, loss: 0.0021863807924091816
step: 280, loss: 0.0013269663322716951
step: 290, loss: 0.0016096904873847961
step: 300, loss: 0.00039003530400805175
step: 310, loss: 0.00023688173678237945
step: 320, loss: 0.001681561698205769
step: 330, loss: 0.00017539146938361228
step: 340, loss: 0.006272868253290653
step: 350, loss: 0.16035933792591095
step: 360, loss: 0.15676143765449524
step: 370, loss: 0.03130869194865227
step: 380, loss: 0.00048802015953697264
step: 390, loss: 0.002388574182987213
step: 400, loss: 0.0013381100725382566
step: 410, loss: 0.00045016835792921484
step: 420, loss: 0.000929532281588763
step: 430, loss: 0.001414925092831254
step: 440, loss: 0.1506846845149994
step: 450, loss: 0.00038956088246777654
step: 460, loss: 0.0140066621825099
step: 470, loss: 0.0581209659576416
step: 480, loss: 0.02073209173977375
step: 490, loss: 0.05863698199391365
step: 500, loss: 0.004073070827871561
step: 510, loss: 0.0011742048664018512
step: 520, loss: 0.00021119067969266325
step: 530, loss: 0.0003438607382122427
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 8: dev_f1=0.9529520295202951, f1=0.9449035812672175,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015321736282203346
step: 10, loss: 0.010195122100412846

```


step: 20, loss: 0.00031197897624224424
step: 30, loss: 0.10260340571403503
step: 40, loss: 0.011136600747704506
step: 50, loss: 0.0004911174182780087
step: 60, loss: 0.0003707358264364302
step: 70, loss: 0.00677760923281312
step: 80, loss: 0.02333776280283928
step: 90, loss: 0.043255362659692764
step: 100, loss: 0.0024759420193731785
step: 110, loss: 0.001053834450431168
step: 120, loss: 0.0001771847455529496
step: 130, loss: 0.00022386584896594286
step: 140, loss: 0.0005194044206291437
step: 150, loss: 0.00014395326434168965
step: 160, loss: 0.00014658019063062966
step: 170, loss: 0.0007456060848198831
step: 180, loss: 0.009347977116703987
step: 190, loss: 0.00011346421524649486
step: 200, loss: 0.0002477254893165082
step: 210, loss: 0.00032418480259366333
step: 220, loss: 0.0028038041200488806
step: 230, loss: 0.0016230514738708735
step: 240, loss: 0.00029448053101077676
step: 250, loss: 0.0009236098849214613
step: 260, loss: 0.0004005772352684289
step: 270, loss: 0.000298697326797992
step: 280, loss: 0.0037410962395370007
step: 290, loss: 0.0005466238944791257
step: 300, loss: 0.000680725381243974
step: 310, loss: 0.00046267054858617485
step: 320, loss: 0.0003882665478158742
step: 330, loss: 0.00038626801688224077
step: 340, loss: 0.0017946816515177488
step: 350, loss: 0.008135910145938396
step: 360, loss: 0.0002619096776470542
step: 370, loss: 0.00019221736874897033
step: 380, loss: 0.0003470016235951334
step: 390, loss: 0.00015391690249089152
step: 400, loss: 0.0010656120721250772
step: 410, loss: 0.014193005859851837
step: 420, loss: 0.00043055255082435906
step: 430, loss: 0.045255452394485474
step: 440, loss: 0.00032114452915266156
step: 450, loss: 0.029427796602249146
step: 460, loss: 0.0008106188615784049
step: 470, loss: 0.00024935073452070355
step: 480, loss: 0.00022233297931961715
step: 490, loss: 0.0024547777138650417

```

step: 500, loss: 0.0001959149813046679
step: 510, loss: 0.0006561438785865903
step: 520, loss: 0.0006905903574079275
step: 530, loss: 0.0004894996527582407
/usr/local/lib/python3.7/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=0.9513108614232209, f1=0.94362292051756,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005257448647171259
step: 10, loss: 0.00017792712606023997
step: 20, loss: 0.028441064059734344
step: 30, loss: 0.04860107973217964
step: 40, loss: 0.0010047531686723232
step: 50, loss: 0.0026639089919626713
step: 60, loss: 0.0004163901030551642
step: 70, loss: 0.00026497122598811984
step: 80, loss: 0.00010231205669697374
step: 90, loss: 8.292988786706701e-05
step: 100, loss: 0.004433069843798876
step: 110, loss: 0.006063737440854311
step: 120, loss: 0.0003871613007504493
step: 130, loss: 0.00015532526595052332
step: 140, loss: 0.001646165968850255
step: 150, loss: 0.0003630288119893521
step: 160, loss: 0.015728147700428963
step: 170, loss: 0.0004140477394685149
step: 180, loss: 0.0019713938236236572
step: 190, loss: 0.0010983687825500965
step: 200, loss: 0.00023938626691233367
step: 210, loss: 0.002827943069860339
step: 220, loss: 0.0004417485324665904
step: 230, loss: 0.00019335892284289002
step: 240, loss: 0.0009958534501492977
step: 250, loss: 0.0002375789190409705
step: 260, loss: 0.0002491497143637389
step: 270, loss: 0.00042656046571210027
step: 280, loss: 0.0072587220929563046
step: 290, loss: 0.16453996300697327
step: 300, loss: 0.007462810724973679
step: 310, loss: 0.024117112159729004
step: 320, loss: 0.015263273380696774
step: 330, loss: 0.007784866262227297
step: 340, loss: 0.005190574564039707
step: 350, loss: 0.0011167565826326609

```

```

step: 360, loss: 0.0008452704059891403
step: 370, loss: 0.002407726366072893
step: 380, loss: 0.00039386283606290817
step: 390, loss: 0.00024872738867998123
step: 400, loss: 0.012371588498353958
step: 410, loss: 0.00024567070067860186
step: 420, loss: 0.0002336471516173333
step: 430, loss: 0.00010122215462615713
step: 440, loss: 0.0001548536674818024
step: 450, loss: 0.0002496932283975184
step: 460, loss: 0.00020264704653527588
step: 470, loss: 0.0021086963824927807
step: 480, loss: 0.00019130932923872024
step: 490, loss: 0.0005989064229652286
step: 500, loss: 0.00039872515480965376
step: 510, loss: 0.00022191926836967468
step: 520, loss: 0.00037107011303305626
step: 530, loss: 0.00020327605307102203
/usr/local/lib/python3.7/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=0.9515377446411928, f1=0.9467838963442852,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.899051656248048e-05
step: 10, loss: 0.0002128138585248962
step: 20, loss: 0.0003296963986940682
step: 30, loss: 0.0006673757452517748
step: 40, loss: 8.078716928139329e-05
step: 50, loss: 0.00022885785438120365
step: 60, loss: 0.00023981480626389384
step: 70, loss: 0.00028028336237184703
step: 80, loss: 0.0004480290226638317
step: 90, loss: 0.0001904438395285979
step: 100, loss: 0.00013457327440846711
step: 110, loss: 0.000599169114138931
step: 120, loss: 0.0001244683371623978
step: 130, loss: 0.004725103732198477
step: 140, loss: 0.00018756510689854622
step: 150, loss: 0.015887422487139702
step: 160, loss: 0.0004622389969881624
step: 170, loss: 0.00024223192303907126
step: 180, loss: 0.00018527318025007844
step: 190, loss: 0.0005615336121991277
step: 200, loss: 0.0003335686633363366
step: 210, loss: 0.000407426618039608

```

```

step: 220, loss: 0.010393696837127209
step: 230, loss: 7.808041118551046e-05
step: 240, loss: 0.001986174378544092
step: 250, loss: 0.0005964107112959027
step: 260, loss: 0.0029128578025847673
step: 270, loss: 0.001677835127338767
step: 280, loss: 0.0004334530094638467
step: 290, loss: 0.017115142196416855
step: 300, loss: 9.491757373325527e-05
step: 310, loss: 0.0566975399851799
step: 320, loss: 0.0001511174050392583
step: 330, loss: 8.420331141678616e-05
step: 340, loss: 0.0010073307203128934
step: 350, loss: 0.00017892058531288058
step: 360, loss: 0.001113148988224566
step: 370, loss: 0.0001596638176124543
step: 380, loss: 0.0004270879435352981
step: 390, loss: 0.0026814125012606382
step: 400, loss: 0.0006100072059780359
step: 410, loss: 0.00011553607328096405
step: 420, loss: 0.00010300763824488968
step: 430, loss: 0.0028488533571362495
step: 440, loss: 0.00029173551592975855
step: 450, loss: 0.00018453219672665
step: 460, loss: 0.00030211234115995467
step: 470, loss: 0.0002772630541585386
step: 480, loss: 0.00023453460016753525
step: 490, loss: 9.486117050983012e-05
step: 500, loss: 0.00023357037571258843
step: 510, loss: 0.001741234795190394
step: 520, loss: 0.0002366711851209402
step: 530, loss: 0.00048243458149954677
/usr/local/lib/python3.7/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=0.9510945505356311, f1=0.9410138248847926,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022328260820358992
step: 10, loss: 0.0001952873426489532
step: 20, loss: 0.0005065201548859477
step: 30, loss: 5.726610106648877e-05
step: 40, loss: 0.0001706702314550057
step: 50, loss: 4.250567508279346e-05
step: 60, loss: 0.00010430687689222395
step: 70, loss: 0.00011493475903989747

```

```
step: 80, loss: 0.2553578317165375
step: 90, loss: 0.00020978125394321978
step: 100, loss: 0.23565180599689484
step: 110, loss: 0.0006461735465563834
step: 120, loss: 0.001164067187346518
step: 130, loss: 0.0004917736514471471
step: 140, loss: 0.000387245265301317
step: 150, loss: 0.00023530620092060417
step: 160, loss: 0.0023570884950459003
step: 170, loss: 0.00026533790514804423
step: 180, loss: 0.00015632054419256747
step: 190, loss: 0.00021913819364272058
step: 200, loss: 0.0014204384060576558
step: 210, loss: 0.00043912691762670875
step: 220, loss: 0.0002211468236055225
step: 230, loss: 7.98197797848843e-05
step: 240, loss: 5.5680844525340945e-05
step: 250, loss: 0.0001662864669924602
step: 260, loss: 8.879784581949934e-05
step: 270, loss: 0.001151294563896954
step: 280, loss: 0.00016872401465661824
step: 290, loss: 8.369003626285121e-05
step: 300, loss: 0.00017627030319999903
step: 310, loss: 0.0008357985061593354
step: 320, loss: 0.00020333037537056953
step: 330, loss: 0.00013915398449171335
step: 340, loss: 0.0009693181491456926
step: 350, loss: 7.396198634523898e-05
step: 360, loss: 7.773334800731391e-05
step: 370, loss: 3.240922524128109e-05
step: 380, loss: 9.952626714948565e-05
step: 390, loss: 0.000230066740186885
step: 400, loss: 4.1795319702941924e-05
step: 410, loss: 0.00023462524404749274
step: 420, loss: 0.00012798949319403619
step: 430, loss: 5.723050344386138e-05
step: 440, loss: 6.43450184725225e-05
step: 450, loss: 0.0008431648020632565
step: 460, loss: 7.205960719147697e-05
step: 470, loss: 0.0015581249026581645
step: 480, loss: 7.71823906688951e-05
step: 490, loss: 0.0001839589240262285
step: 500, loss: 7.288864435395226e-05
step: 510, loss: 0.00024112437677104026
step: 520, loss: 0.0003062636242248118
step: 530, loss: 0.00013384639169089496
/usr/local/lib/python3.7/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=0.9499770536943551, f1=0.943155979990905,
best_f1=0.937471051412691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.906753747491166e-05
step: 10, loss: 9.053220128407702e-05
step: 20, loss: 0.00026370803243480623
step: 30, loss: 0.00025466986699029803
step: 40, loss: 0.0005303655634634197
step: 50, loss: 0.004150615073740482
step: 60, loss: 0.00010663560533430427
step: 70, loss: 0.00019544271344784647
step: 80, loss: 0.0022593399044126272
step: 90, loss: 0.007328339386731386
step: 100, loss: 0.00010651012416929007
step: 110, loss: 0.0005767618422396481
step: 120, loss: 9.703163959784433e-05
step: 130, loss: 4.592115874402225e-05
step: 140, loss: 0.001314118504524231
step: 150, loss: 0.00379954744130373
step: 160, loss: 0.0003206537803635001
step: 170, loss: 7.770696538500488e-05
step: 180, loss: 0.0002490339975338429
step: 190, loss: 0.00036187091609463096
step: 200, loss: 0.00026418865309096873
step: 210, loss: 0.00023525397409684956
step: 220, loss: 0.00015377603995148093
step: 230, loss: 0.0015033312374725938
step: 240, loss: 0.0002600644656922668
step: 250, loss: 0.00020354420121293515
step: 260, loss: 0.00015920605801511556
step: 270, loss: 0.008385665714740753
step: 280, loss: 0.00021666003158316016
step: 290, loss: 5.5210017308127135e-05
step: 300, loss: 0.00016664979921188205
step: 310, loss: 0.0001361674367217347
step: 320, loss: 9.962861076928675e-05
step: 330, loss: 0.0006474130786955357
step: 340, loss: 0.00021262463997118175
step: 350, loss: 4.29696447099559e-05
step: 360, loss: 0.0008871880709193647
step: 370, loss: 0.00016584982222411782
step: 380, loss: 0.00030843273270875216
step: 390, loss: 9.058241994353011e-05
step: 400, loss: 0.1688692569732666
step: 410, loss: 0.00023072048497851938
```

```

step: 420, loss: 0.00020946480799466372
step: 430, loss: 0.0003949782403651625
step: 440, loss: 0.0001259086129721254
step: 450, loss: 0.00035865293466486037
step: 460, loss: 0.0005508443573489785
step: 470, loss: 7.165165152400732e-05
step: 480, loss: 8.772669389145449e-05
step: 490, loss: 0.005035775247961283
step: 500, loss: 0.0003313605848234147
step: 510, loss: 0.00012061512825312093
step: 520, loss: 0.00017297915474046022
step: 530, loss: 0.00016167784633580595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9547363509099394, f1=0.9420423183072677,
best_f1=0.9420423183072677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014110974734649062
step: 10, loss: 0.00010939583444269374
step: 20, loss: 9.367047459818423e-05
step: 30, loss: 0.00021086666674818844
step: 40, loss: 0.00011612067464739084
step: 50, loss: 0.0006683355895802379
step: 60, loss: 0.00014850342995487154
step: 70, loss: 0.00017013491014949977
step: 80, loss: 0.000726998143363744
step: 90, loss: 7.764819747535512e-05
step: 100, loss: 0.00012493194662965834
step: 110, loss: 0.001046891906298697
step: 120, loss: 0.00012022001465084031
step: 130, loss: 0.00012635112216230482
step: 140, loss: 0.00014671566896140575
step: 150, loss: 0.00021690675930585712
step: 160, loss: 0.0022594102192670107
step: 170, loss: 0.0012987875379621983
step: 180, loss: 7.535974873462692e-05
step: 190, loss: 0.00014988754992373288
step: 200, loss: 0.00014510590699501336
step: 210, loss: 0.00020430097356438637
step: 220, loss: 0.00013495111488737166
step: 230, loss: 0.0002017575898207724

```

```

step: 240, loss: 6.604770169360563e-05
step: 250, loss: 0.00013009223039261997
step: 260, loss: 8.088299364317209e-05
step: 270, loss: 0.0006540139438584447
step: 280, loss: 0.0004115173942409456
step: 290, loss: 8.624742622487247e-05
step: 300, loss: 0.00012966380745638162
step: 310, loss: 0.0006200367934070528
step: 320, loss: 0.00011800337233580649
step: 330, loss: 0.0001896908215712756
step: 340, loss: 0.00019763234013225883
step: 350, loss: 0.0002045456931227818
step: 360, loss: 9.560545004205778e-05
step: 370, loss: 0.000734862987883389
step: 380, loss: 0.0020767792593687773
step: 390, loss: 0.00022821887978352606
step: 400, loss: 0.0009123578784056008
step: 410, loss: 5.919201794313267e-05
step: 420, loss: 7.17452639946714e-05
step: 430, loss: 0.0032208317425101995
step: 440, loss: 0.00015499941946472973
step: 450, loss: 8.23301452328451e-05
step: 460, loss: 0.00029046344570815563
step: 470, loss: 0.0003038433496840298
step: 480, loss: 7.592151814606041e-05
step: 490, loss: 0.00011278921738266945
step: 500, loss: 7.343234756262973e-05
step: 510, loss: 0.00043051462853327394
step: 520, loss: 9.298031363869086e-05
step: 530, loss: 0.0001031628780765459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9558755225267069, f1=0.9459211732355638,
best_f1=0.9459211732355638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012267952843103558
step: 10, loss: 0.00012974954734090716
step: 20, loss: 0.0001219552505062893
step: 30, loss: 0.00014848136925138533
step: 40, loss: 9.888113709166646e-05
step: 50, loss: 0.00014150413335300982

```


step: 60, loss: 0.00019528286065906286
step: 70, loss: 6.926352216396481e-05
step: 80, loss: 0.00011182337038917467
step: 90, loss: 0.0001715514954412356
step: 100, loss: 5.269592293188907e-05
step: 110, loss: 0.0002491829509381205
step: 120, loss: 9.247918205801398e-05
step: 130, loss: 0.00014332339924294502
step: 140, loss: 0.027166124433279037
step: 150, loss: 0.00013147233403287828
step: 160, loss: 0.00017703042249195278
step: 170, loss: 4.507733683567494e-05
step: 180, loss: 0.00021049987117294222
step: 190, loss: 0.0009522888576611876
step: 200, loss: 8.856225031195208e-05
step: 210, loss: 5.640763993142173e-05
step: 220, loss: 8.539736154489219e-05
step: 230, loss: 0.20548629760742188
step: 240, loss: 0.00015324768901336938
step: 250, loss: 0.00014842100790701807
step: 260, loss: 7.363538315985352e-05
step: 270, loss: 0.0002718755276873708
step: 280, loss: 6.231716542970389e-05
step: 290, loss: 0.0001281712029594928
step: 300, loss: 6.468347419286147e-05
step: 310, loss: 0.024950696155428886
step: 320, loss: 0.0003876594128087163
step: 330, loss: 0.00027160230092704296
step: 340, loss: 0.00012514293484855443
step: 350, loss: 0.014711800962686539
step: 360, loss: 0.0006916792481206357
step: 370, loss: 0.001760148094035685
step: 380, loss: 0.00015638793411199003
step: 390, loss: 0.00012762202823068947
step: 400, loss: 0.00044084584806114435
step: 410, loss: 0.00010259871487505734
step: 420, loss: 9.810437040869147e-05
step: 430, loss: 3.633803135016933e-05
step: 440, loss: 0.00022017580340616405
step: 450, loss: 0.0012324348790571094
step: 460, loss: 8.139191777445376e-05
step: 470, loss: 8.237801375798881e-05
step: 480, loss: 8.929391333367676e-05
step: 490, loss: 0.00016504449013154954
step: 500, loss: 0.00012455755495466292
step: 510, loss: 7.526592526119202e-05
step: 520, loss: 0.0009575159638188779
step: 530, loss: 5.387249620980583e-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=0.9564007421150279, f1=0.9461187214611873,
best_f1=0.9461187214611873
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.")
5742it [00:23, 248.67it/s]
load_f1 = 0.9551548774849746
real_f1 = 0.9559981472904122
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 [00:22, 199.52it/s]

```

2.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4995102882385254
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41936203837394714
step: 20, loss: 0.4709685444831848
step: 30, loss: 0.368865042924881
step: 40, loss: 0.30103373527526855
step: 50, loss: 0.43656909465789795
step: 60, loss: 0.5244560837745667
step: 70, loss: 0.30361050367355347
step: 80, loss: 0.35396912693977356
step: 90, loss: 0.2150481939315796
step: 100, loss: 0.2782241106033325
step: 110, loss: 0.246608167886734
step: 120, loss: 0.38572368025779724
step: 130, loss: 0.24525296688079834
step: 140, loss: 0.5049907565116882
step: 150, loss: 0.4160477817058563
step: 160, loss: 0.5327482223510742
step: 170, loss: 0.24319611489772797
step: 180, loss: 0.3899673521518707
step: 190, loss: 0.6602665185928345
step: 200, loss: 0.39211663603782654
step: 210, loss: 0.5099440217018127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36444979906082153
step: 10, loss: 0.19533996284008026
step: 20, loss: 0.49549779295921326
step: 30, loss: 0.4966703951358795
step: 40, loss: 0.4898938238620758
step: 50, loss: 0.24358461797237396
step: 60, loss: 0.3118874132633209
step: 70, loss: 0.43183842301368713
step: 80, loss: 0.3270207941532135
step: 90, loss: 0.3623496890068054
step: 100, loss: 0.5043300986289978
step: 110, loss: 0.3784053325653076
step: 120, loss: 0.24790897965431213
step: 130, loss: 0.16632522642612457
step: 140, loss: 0.23783735930919647
step: 150, loss: 0.4438115656375885
step: 160, loss: 0.16078557074069977
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.6344438195228577
step: 180, loss: 0.31634587049484253
step: 190, loss: 0.3126146197319031
step: 200, loss: 0.1499144434928894
step: 210, loss: 0.3079928457736969
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2548840641975403
step: 10, loss: 0.24764512479305267
step: 20, loss: 0.4828086793422699
step: 30, loss: 0.2510899305343628
step: 40, loss: 0.42296120524406433
step: 50, loss: 0.48696401715278625
step: 60, loss: 0.5139468908309937
step: 70, loss: 0.20191670954227448
step: 80, loss: 0.46969977021217346
step: 90, loss: 0.24354660511016846
step: 100, loss: 0.361467570066452

```

```

step: 110, loss: 0.2551610469818115
step: 120, loss: 0.2326686829328537
step: 130, loss: 0.163332998752594
step: 140, loss: 0.371633917093277
step: 150, loss: 0.31219199299812317
step: 160, loss: 0.21446363627910614
step: 170, loss: 0.3763662874698639
step: 180, loss: 0.2513311505317688
step: 190, loss: 0.16173627972602844
step: 200, loss: 0.22179844975471497
step: 210, loss: 0.2627320885658264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3014991879463196
step: 10, loss: 0.3036693334579468
step: 20, loss: 0.30517247319221497
step: 30, loss: 0.23620961606502533
step: 40, loss: 0.23634116351604462
step: 50, loss: 0.2406916469335556
step: 60, loss: 0.46619683504104614
step: 70, loss: 0.21195046603679657
step: 80, loss: 0.2415606826543808
step: 90, loss: 0.5272196531295776
step: 100, loss: 0.3805077373981476
step: 110, loss: 0.6020689010620117
step: 120, loss: 0.3812374174594879
step: 130, loss: 0.6691019535064697
step: 140, loss: 0.5075435638427734
step: 150, loss: 0.3835507333278656
step: 160, loss: 0.31339359283447266
step: 170, loss: 0.15909482538700104
step: 180, loss: 0.10089802742004395
step: 190, loss: 0.18395967781543732
step: 200, loss: 0.3097943365573883
step: 210, loss: 0.41269853711128235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.29914529914529914, f1=0.3205268935236005,
best_f1=0.3205268935236005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35541489720344543
step: 10, loss: 0.285169392824173
step: 20, loss: 0.31854748725891113
step: 30, loss: 0.2607403099536896
step: 40, loss: 0.4293382465839386
step: 50, loss: 0.38837161660194397
step: 60, loss: 0.38887879252433777
step: 70, loss: 0.2497842013835907
step: 80, loss: 0.44314491748809814
step: 90, loss: 0.44395673274993896
step: 100, loss: 0.24457217752933502
step: 110, loss: 0.1679259091615677
step: 120, loss: 0.24292796850204468
step: 130, loss: 0.3160398304462433
step: 140, loss: 0.5364706516265869
step: 150, loss: 0.31767114996910095
step: 160, loss: 0.2462482750415802
step: 170, loss: 0.3767916262149811
step: 180, loss: 0.23944544792175293
step: 190, loss: 0.5372313857078552
step: 200, loss: 0.37742140889167786
step: 210, loss: 0.24213773012161255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.3205268935236005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15820656716823578
step: 10, loss: 0.37639400362968445
step: 20, loss: 0.3074698746204376
step: 30, loss: 0.23806248605251312
step: 40, loss: 0.3158939480781555
step: 50, loss: 0.5889603495597839
step: 60, loss: 0.2530084252357483
step: 70, loss: 0.3294301927089691
step: 80, loss: 0.2585921287536621
step: 90, loss: 0.3011522889137268
step: 100, loss: 0.2687426805496216
step: 110, loss: 0.40076446533203125
step: 120, loss: 0.24193838238716125

```

```

step: 130, loss: 0.394644558429718
step: 140, loss: 0.38952669501304626
step: 150, loss: 0.17599712312221527
step: 160, loss: 0.15796850621700287
step: 170, loss: 0.4401673674583435
step: 180, loss: 0.3748610317707062
step: 190, loss: 0.4520539343357086
step: 200, loss: 0.3126935064792633
step: 210, loss: 0.3829006552696228
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.3205268935236005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4408193528652191
step: 10, loss: 0.3078269362449646
step: 20, loss: 0.4528360962867737
step: 30, loss: 0.2471923977136612
step: 40, loss: 0.2625996470451355
step: 50, loss: 0.37660136818885803
step: 60, loss: 0.31459251046180725
step: 70, loss: 0.24020931124687195
step: 80, loss: 0.4278430938720703
step: 90, loss: 0.37075555324554443
step: 100, loss: 0.3936060667037964
step: 110, loss: 0.3175524175167084
step: 120, loss: 0.300905704498291
step: 130, loss: 0.38580572605133057
step: 140, loss: 0.31281906366348267
step: 150, loss: 0.31959566473960876
step: 160, loss: 0.568099856376648
step: 170, loss: 0.556435227394104
step: 180, loss: 0.22604168951511383
step: 190, loss: 0.2578217387199402
step: 200, loss: 0.28849390149116516
step: 210, loss: 0.30857157707214355
/usr/local/lib/python3.7/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=0.2813559322033898, f1=0.315512708150745,
best_f1=0.3205268935236005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5280846953392029

```



```

step: 10, loss: 0.30813515186309814
step: 20, loss: 0.20374515652656555
step: 30, loss: 0.25878241658210754
step: 40, loss: 0.2424544095993042
step: 50, loss: 0.13836391270160675
step: 60, loss: 0.18094374239444733
step: 70, loss: 0.4296761155128479
step: 80, loss: 0.2795296013355255
step: 90, loss: 0.38256514072418213
step: 100, loss: 0.5190650224685669
step: 110, loss: 0.2906706631183624
step: 120, loss: 0.32166633009910583
step: 130, loss: 0.13436760008335114
step: 140, loss: 0.31794437766075134
step: 150, loss: 0.5034316778182983
step: 160, loss: 0.44179680943489075
step: 170, loss: 0.528410792350769
step: 180, loss: 0.24461661279201508
step: 190, loss: 0.20047231018543243
step: 200, loss: 0.3766734302043915
step: 210, loss: 0.3898133635520935
/usr/local/lib/python3.7/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=0.2985074626865672, f1=0.33531510107015455,
best_f1=0.3205268935236005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.47716283798217773
step: 10, loss: 0.2634734511375427
step: 20, loss: 0.34742823243141174
step: 30, loss: 0.13498613238334656
step: 40, loss: 0.17965252697467804
step: 50, loss: 0.41380852460861206
step: 60, loss: 0.11664088815450668
step: 70, loss: 0.32733118534088135
step: 80, loss: 0.28139933943748474
step: 90, loss: 0.26912909746170044
step: 100, loss: 0.4587562382221222
step: 110, loss: 0.4922640919685364
step: 120, loss: 0.38103222846984863
step: 130, loss: 0.1734798401594162
step: 140, loss: 0.6107993125915527
step: 150, loss: 0.17649444937705994
step: 160, loss: 0.2715303897857666
step: 170, loss: 0.2921255826950073
step: 180, loss: 0.45270586013793945

```

```

step: 190, loss: 0.2134733349084854
step: 200, loss: 0.21891266107559204
step: 210, loss: 0.296622097492218
/usr/local/lib/python3.7/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=0.29629629629629634, f1=0.2918454935622318,
best_f1=0.3205268935236005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27094167470932007
step: 10, loss: 0.40532785654067993
step: 20, loss: 0.22961057722568512
step: 30, loss: 0.1589590311050415
step: 40, loss: 0.30144983530044556
step: 50, loss: 0.43729934096336365
step: 60, loss: 0.1891823709011078
step: 70, loss: 0.4085310101509094
step: 80, loss: 0.1754636913537979
step: 90, loss: 0.37627893686294556
step: 100, loss: 0.44846591353416443
step: 110, loss: 0.12636995315551758
step: 120, loss: 0.45838844776153564
step: 130, loss: 0.2803381681442261
step: 140, loss: 0.2525870203971863
step: 150, loss: 0.30632269382476807
step: 160, loss: 0.2283470779657364
step: 170, loss: 0.21074256300926208
step: 180, loss: 0.2832314074039459
step: 190, loss: 0.19395595788955688
step: 200, loss: 0.5233060121536255
step: 210, loss: 0.2219059318304062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.31812255541069095, f1=0.3250975292587776,
best_f1=0.3250975292587776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25794392824172974
step: 10, loss: 0.2595300078392029
step: 20, loss: 0.34813550114631653

```

```

step: 30, loss: 0.1609373688697815
step: 40, loss: 0.35499322414398193
step: 50, loss: 0.37429186701774597
step: 60, loss: 0.2647121846675873
step: 70, loss: 0.16207659244537354
step: 80, loss: 0.34314578771591187
step: 90, loss: 0.33234211802482605
step: 100, loss: 0.37879979610443115
step: 110, loss: 0.39745354652404785
step: 120, loss: 0.3378097712993622
step: 130, loss: 0.2024555802345276
step: 140, loss: 0.3173360824584961
step: 150, loss: 0.2903156578540802
step: 160, loss: 0.12941306829452515
step: 170, loss: 0.31591683626174927
step: 180, loss: 0.25471678376197815
step: 190, loss: 0.39088520407676697
step: 200, loss: 0.16922518610954285
step: 210, loss: 0.3858330547809601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3310580204778157, f1=0.3382594417077176,
best_f1=0.3382594417077176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21692602336406708
step: 10, loss: 0.29362234473228455
step: 20, loss: 0.39642611145973206
step: 30, loss: 0.31075820326805115
step: 40, loss: 0.18493331968784332
step: 50, loss: 0.41098618507385254
step: 60, loss: 0.09808807075023651
step: 70, loss: 0.3681352138519287
step: 80, loss: 0.3255402445793152
step: 90, loss: 0.37224724888801575
step: 100, loss: 0.11340014636516571
step: 110, loss: 0.18427041172981262
step: 120, loss: 0.1553625911474228
step: 130, loss: 0.4626527428627014
step: 140, loss: 0.4540068209171295
step: 150, loss: 0.10825157165527344
step: 160, loss: 0.18298687040805817

```

```

step: 170, loss: 0.15501143038272858
step: 180, loss: 0.3428415358066559
step: 190, loss: 0.26330605149269104
step: 200, loss: 0.11694702506065369
step: 210, loss: 0.26681986451148987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.33384853168469864, f1=0.38562091503267976,
best_f1=0.38562091503267976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2249579280614853
step: 10, loss: 0.14948879182338715
step: 20, loss: 0.1778249740600586
step: 30, loss: 0.14079345762729645
step: 40, loss: 0.09636328369379044
step: 50, loss: 0.20881909132003784
step: 60, loss: 0.3550778925418854
step: 70, loss: 0.40222424268722534
step: 80, loss: 0.3160936236381531
step: 90, loss: 0.2530174255371094
step: 100, loss: 0.3398857116699219
step: 110, loss: 0.1806141436100006
step: 120, loss: 0.17222368717193604
step: 130, loss: 0.21109379827976227
step: 140, loss: 0.21872591972351074
step: 150, loss: 0.29376357793807983
step: 160, loss: 0.2823300063610077
step: 170, loss: 0.2001647800207138
step: 180, loss: 0.0957135334610939
step: 190, loss: 0.24325774610042572
step: 200, loss: 0.36115145683288574
step: 210, loss: 0.2965213358402252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37322515212981744, f1=0.40591966173361516,
best_f1=0.40591966173361516

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20067527890205383
step: 10, loss: 0.28340473771095276
step: 20, loss: 0.24646353721618652
step: 30, loss: 0.14372670650482178
step: 40, loss: 0.28143012523651123
step: 50, loss: 0.26013290882110596
step: 60, loss: 0.4005483090877533
step: 70, loss: 0.23454301059246063
step: 80, loss: 0.2219592034816742
step: 90, loss: 0.11092588305473328
step: 100, loss: 0.11964017152786255
step: 110, loss: 0.37826451659202576
step: 120, loss: 0.15081557631492615
step: 130, loss: 0.1625194251537323
step: 140, loss: 0.23524145781993866
step: 150, loss: 0.27838975191116333
step: 160, loss: 0.15054388344287872
step: 170, loss: 0.2138317972421646
step: 180, loss: 0.15663009881973267
step: 190, loss: 0.3708851635456085
step: 200, loss: 0.06592334061861038
step: 210, loss: 0.26570799946784973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.39672801635991817, f1=0.42241379310344823,
best_f1=0.42241379310344823
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19061337411403656
step: 10, loss: 0.22724813222885132
step: 20, loss: 0.3511189818382263
step: 30, loss: 0.15060803294181824
step: 40, loss: 0.40424561500549316
step: 50, loss: 0.09451291710138321
step: 60, loss: 0.2685280442237854
step: 70, loss: 0.3828473389148712
step: 80, loss: 0.23960204422473907
step: 90, loss: 0.2961353063583374
step: 100, loss: 0.19626867771148682

```

```

step: 110, loss: 0.19109341502189636
step: 120, loss: 0.2338823825120926
step: 130, loss: 0.1973849982023239
step: 140, loss: 0.2829414904117584
step: 150, loss: 0.30466482043266296
step: 160, loss: 0.3019488453865051
step: 170, loss: 0.2370266318321228
step: 180, loss: 0.2393067181110382
step: 190, loss: 0.1974155753850937
step: 200, loss: 0.22947697341442108
step: 210, loss: 0.3384624421596527
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4000000000000001, f1=0.44705882352941173,
best_f1=0.44705882352941173
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 [00:05, 393.80it/s]
load_f1 = 0.40322580645161293
real_f1 = 0.3834422657952069
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 201.42it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

Selected optimization level 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: 0.450461208820343
step: 10, loss: 0.37340807914733887
step: 20, loss: 0.26923197507858276
step: 30, loss: 0.4495473802089691
step: 40, loss: 0.2613702714443207
step: 50, loss: 0.30123305320739746
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.5271520018577576
step: 70, loss: 0.4617880582809448
step: 80, loss: 0.16493107378482819
step: 90, loss: 0.3588343560695648
step: 100, loss: 0.47342416644096375
step: 110, loss: 0.2366301268339157
step: 120, loss: 0.3486109673976898
step: 130, loss: 0.35191214084625244
step: 140, loss: 0.15988977253437042
step: 150, loss: 0.30743297934532166
step: 160, loss: 0.25709694623947144

```



```

step: 170, loss: 0.3933044373989105
step: 180, loss: 0.15996979176998138
step: 190, loss: 0.1611507385969162
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4112624526023865
step: 10, loss: 0.3040861487388611
step: 20, loss: 0.6118457913398743
step: 30, loss: 0.2368740290403366
step: 40, loss: 0.5299840569496155
step: 50, loss: 0.3569798469543457
step: 60, loss: 0.45612549781799316
step: 70, loss: 0.30345627665519714
step: 80, loss: 0.15923789143562317
step: 90, loss: 0.33646920323371887
step: 100, loss: 0.2549055814743042
step: 110, loss: 0.360750675201416
step: 120, loss: 0.24302273988723755
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5503145456314087
step: 140, loss: 0.3266708552837372
step: 150, loss: 0.3249206840991974
step: 160, loss: 0.2959466874599457
step: 170, loss: 0.24326463043689728
step: 180, loss: 0.17538313567638397
step: 190, loss: 0.23364830017089844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37515345215797424
step: 10, loss: 0.400308758020401
step: 20, loss: 0.4691533148288727
step: 30, loss: 0.2753675878047943

```

```

step: 40, loss: 0.0777311697602272
step: 50, loss: 0.3291734755039215
step: 60, loss: 0.11361335963010788
step: 70, loss: 0.2904478907585144
step: 80, loss: 0.29618725180625916
step: 90, loss: 0.3657459616661072
step: 100, loss: 0.6261195540428162
step: 110, loss: 0.6053341627120972
step: 120, loss: 0.3855575919151306
step: 130, loss: 0.1274072527885437
step: 140, loss: 0.17609722912311554
step: 150, loss: 0.4453984797000885
step: 160, loss: 0.2349245697259903
step: 170, loss: 0.3108481168746948
step: 180, loss: 0.3704390525817871
step: 190, loss: 0.13460186123847961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7179487179487181, f1=0.7572254335260116,
best_f1=0.7572254335260116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08434134721755981
step: 10, loss: 0.2121339589357376
step: 20, loss: 0.18632948398590088
step: 30, loss: 0.042264699935913086
step: 40, loss: 0.061929576098918915
step: 50, loss: 0.15660139918327332
step: 60, loss: 0.1345098614692688
step: 70, loss: 0.046613212674856186
step: 80, loss: 0.01656983606517315
step: 90, loss: 0.044480882585048676
step: 100, loss: 0.06002200022339821
step: 110, loss: 0.2833907902240753
step: 120, loss: 0.12582260370254517
step: 130, loss: 0.15125779807567596
step: 140, loss: 0.27330633997917175
step: 150, loss: 0.028759021311998367
step: 160, loss: 0.1349562555551529
step: 170, loss: 0.1083955392241478
step: 180, loss: 0.091163270175457
step: 190, loss: 0.026525462046265602

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8057142857142858, f1=0.8057142857142858,
best_f1=0.8057142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06932180374860764
step: 10, loss: 0.11789186298847198
step: 20, loss: 0.06330090761184692
step: 30, loss: 0.09775909036397934
step: 40, loss: 0.056221552193164825
step: 50, loss: 0.13388845324516296
step: 60, loss: 0.010850788094103336
step: 70, loss: 0.1852722018957138
step: 80, loss: 0.0689152181148529
step: 90, loss: 0.16444896161556244
step: 100, loss: 0.10488774627447128
step: 110, loss: 0.01743420958518982
step: 120, loss: 0.029883628711104393
step: 130, loss: 0.25183847546577454
step: 140, loss: 0.13205082714557648
step: 150, loss: 0.14705193042755127
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 160, loss: 0.15870848298072815
step: 170, loss: 0.011575395241379738
step: 180, loss: 0.022358493879437447
step: 190, loss: 0.11652500182390213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8086253369272237, f1=0.8315217391304349,
best_f1=0.8315217391304349
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11005239188671112
step: 10, loss: 0.17062173783779144
step: 20, loss: 0.15275061130523682

```

```

step: 30, loss: 0.13080735504627228
step: 40, loss: 0.017134513705968857
step: 50, loss: 0.1111084371805191
step: 60, loss: 0.1396273374557495
step: 70, loss: 0.17492029070854187
step: 80, loss: 0.028179023414850235
step: 90, loss: 0.01072810310870409
step: 100, loss: 0.1474597454071045
step: 110, loss: 0.02820683643221855
step: 120, loss: 0.03827730938792229
step: 130, loss: 0.17957942187786102
step: 140, loss: 0.03002110682427883
step: 150, loss: 0.07363259792327881
step: 160, loss: 0.16537901759147644
step: 170, loss: 0.25516900420188904
step: 180, loss: 0.025246454402804375
step: 190, loss: 0.21418558061122894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8100558659217876, f1=0.8291316526610644,
best_f1=0.8291316526610644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1732795387506485
step: 10, loss: 0.017411433160305023
step: 20, loss: 0.010201537050306797
step: 30, loss: 0.08069363981485367
step: 40, loss: 0.009109221398830414
step: 50, loss: 0.003997885622084141
step: 60, loss: 0.20155119895935059
step: 70, loss: 0.006766339763998985
step: 80, loss: 0.05077743902802467
step: 90, loss: 0.010033776983618736
step: 100, loss: 0.3027651011943817
step: 110, loss: 0.2819867730140686
step: 120, loss: 0.14300653338432312
step: 130, loss: 0.08394601196050644
step: 140, loss: 0.18279215693473816
step: 150, loss: 0.03215031325817108
step: 160, loss: 0.2387307584285736
step: 170, loss: 0.03363984823226929
step: 180, loss: 0.05746108666062355

```

```

step: 190, loss: 0.106028251349926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.824858757062147, f1=0.8324022346368716,
best_f1=0.8324022346368716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13493919372558594
step: 10, loss: 0.06452923268079758
step: 20, loss: 0.014633027836680412
step: 30, loss: 0.1827663630247116
step: 40, loss: 0.11511289328336716
step: 50, loss: 0.03767140954732895
step: 60, loss: 0.06595633178949356
step: 70, loss: 0.02347422018647194
step: 80, loss: 0.18271909654140472
step: 90, loss: 0.01346944272518158
step: 100, loss: 0.008262630552053452
step: 110, loss: 0.012975755147635937
step: 120, loss: 0.010214942507445812
step: 130, loss: 0.01047370582818985
step: 140, loss: 0.02360975556075573
step: 150, loss: 0.07157906144857407
step: 160, loss: 0.009030315093696117
step: 170, loss: 0.003087081480771303
step: 180, loss: 0.1848897486925125
step: 190, loss: 0.023086490109562874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8272251308900523, f1=0.84375, best_f1=0.84375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026649456471204758
step: 10, loss: 0.0038229324854910374
step: 20, loss: 0.07241310179233551
step: 30, loss: 0.006604415364563465

```

```

step: 40, loss: 0.1829276978969574
step: 50, loss: 0.01302122138440609
step: 60, loss: 0.12388494610786438
step: 70, loss: 0.007056831382215023
step: 80, loss: 0.09561007469892502
step: 90, loss: 0.038421664386987686
step: 100, loss: 0.08258239179849625
step: 110, loss: 0.004364853724837303
step: 120, loss: 0.021374603733420372
step: 130, loss: 0.10240643471479416
step: 140, loss: 0.010897214524447918
step: 150, loss: 0.0631970465183258
step: 160, loss: 0.003806926542893052
step: 170, loss: 0.04133901372551918
step: 180, loss: 0.11456143856048584
step: 190, loss: 0.0028077231254428625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8273972602739726, f1=0.848314606741573,
best_f1=0.848314606741573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00603967672213912
step: 10, loss: 0.0022528881672769785
step: 20, loss: 0.0014071417972445488
step: 30, loss: 0.15527001023292542
step: 40, loss: 0.0038916547782719135
step: 50, loss: 0.034093961119651794
step: 60, loss: 0.006638498045504093
step: 70, loss: 0.0023738257586956024
step: 80, loss: 0.016250526532530785
step: 90, loss: 0.0033156732097268105
step: 100, loss: 0.08814922720193863
step: 110, loss: 0.010810177773237228
step: 120, loss: 0.04109019413590431
step: 130, loss: 0.0031702572014182806
step: 140, loss: 0.0020676287822425365
step: 150, loss: 0.001937143737450242
step: 160, loss: 0.0044590807519853115
step: 170, loss: 0.0033506155014038086
step: 180, loss: 0.011369806714355946
step: 190, loss: 0.00575296301394701

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8421052631578948, f1=0.847457627118644,
best_f1=0.847457627118644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05593300610780716
step: 10, loss: 0.003299335017800331
step: 20, loss: 0.012963848188519478
step: 40, loss: 0.003266336163505912
step: 50, loss: 0.00143732491414994
step: 60, loss: 0.009419720619916916
step: 70, loss: 0.002596953883767128
step: 80, loss: 0.25735604763031006
step: 90, loss: 0.01095598004758358
step: 100, loss: 0.00490774167701602
step: 110, loss: 0.15952825546264648
step: 120, loss: 0.004803864751011133
step: 130, loss: 0.022807978093624115
step: 140, loss: 0.006444510072469711
step: 150, loss: 0.017158659175038338
step: 160, loss: 0.012202556245028973
step: 170, loss: 0.02219068817794323
step: 180, loss: 0.0008304127841256559
step: 190, loss: 0.0018410763004794717
/usr/local/lib/python3.7/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=0.8196721311475409, f1=0.8324022346368716,
best_f1=0.847457627118644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035840426571667194
step: 10, loss: 0.0036041338462382555
step: 20, loss: 0.002844554837793112
step: 30, loss: 0.07098616659641266
step: 40, loss: 0.0018331286264583468
step: 50, loss: 0.030947046354413033
step: 60, loss: 0.006172133609652519
step: 70, loss: 0.00344030256383121
step: 80, loss: 0.004259011242538691

```

```

step: 90, loss: 0.0008407521527260542
step: 100, loss: 0.0014920576941221952
step: 110, loss: 0.14597557485103607
step: 120, loss: 0.03188549354672432
step: 130, loss: 0.007260757032781839
step: 140, loss: 0.005992017686367035
step: 150, loss: 0.004102362785488367
step: 160, loss: 0.0015866841422393918
step: 170, loss: 0.0171247199177742
step: 180, loss: 0.001566300867125392
step: 190, loss: 0.0009993212297558784
/usr/local/lib/python3.7/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=0.8264462809917356, f1=0.8324022346368716,
best_f1=0.847457627118644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06721014529466629
step: 10, loss: 0.0014727518428117037
step: 20, loss: 0.0017319332109764218
step: 30, loss: 0.018374981358647346
step: 40, loss: 0.0055464752949774265
step: 50, loss: 0.011692128144204617
step: 60, loss: 0.0010938465129584074
step: 70, loss: 0.055834125727415085
step: 80, loss: 0.000847881892696023
step: 90, loss: 0.07059399783611298
step: 100, loss: 0.02465212717652321
step: 110, loss: 0.0056112115271389484
step: 120, loss: 0.0006526736542582512
step: 130, loss: 0.013654058799147606
step: 140, loss: 0.009707659482955933
step: 150, loss: 0.0044563692063093185
step: 160, loss: 0.001168272690847516
step: 170, loss: 0.004776609130203724
step: 180, loss: 0.0012696662452071905
step: 190, loss: 0.1184300035238266
/usr/local/lib/python3.7/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=0.8254847645429362, f1=0.836676217765043,
best_f1=0.847457627118644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004231587052345276

```



```

step: 10, loss: 0.0008403437677770853
step: 20, loss: 0.0023005674593150616
step: 30, loss: 0.003680718597024679
step: 40, loss: 0.010632255114614964
step: 50, loss: 0.00551948556676507
step: 60, loss: 0.006188856437802315
step: 70, loss: 0.006294284947216511
step: 80, loss: 0.0021676714532077312
step: 90, loss: 0.0010597719810903072
step: 100, loss: 0.001666775904595852
step: 110, loss: 0.003190354211255908
step: 120, loss: 0.013030105270445347
step: 130, loss: 0.0008731014677323401
step: 140, loss: 0.0008740610210224986
step: 150, loss: 0.0012096988502889872
step: 160, loss: 0.013912429101765156
step: 170, loss: 0.044264547526836395
step: 180, loss: 0.014190842397511005
step: 190, loss: 0.0044016409665346146
/usr/local/lib/python3.7/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=0.8244680851063829, f1=0.8392370572207084,
best_f1=0.847457627118644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029306700453162193
step: 10, loss: 0.0016637961380183697
step: 20, loss: 0.002340614562854171
step: 30, loss: 0.012717712670564651
step: 40, loss: 0.0027614671271294355
step: 50, loss: 0.0023600957356393337
step: 60, loss: 0.0005478534149006009
step: 70, loss: 0.005124840885400772
step: 80, loss: 0.00030341112869791687
step: 90, loss: 0.03164338693022728
step: 100, loss: 0.00034223790862597525
step: 110, loss: 0.00315861776471138
step: 120, loss: 0.0007710672216489911
step: 130, loss: 0.001096124411560595
step: 140, loss: 0.012350893579423428
step: 150, loss: 0.0006350439507514238
step: 160, loss: 0.0009723668918013573
step: 170, loss: 0.005035065114498138
step: 180, loss: 0.0023517932277172804
step: 190, loss: 0.046582482755184174
/usr/local/lib/python3.7/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=0.8257372654155496, f1=0.8328767123287673,
best_f1=0.847457627118644
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 [00:09, 226.75it/s]
load_f1 = 0.837465564738292
real_f1 = 0.8351648351648351
733it [00:00, 3187.50it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.19it/s]

```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4927927255630493
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41033512353897095
step: 20, loss: 0.3902902901172638
step: 30, loss: 0.37784257531166077
step: 40, loss: 0.519510805606842
step: 50, loss: 0.312193900346756
step: 60, loss: 0.5999280214309692
step: 70, loss: 0.33485081791877747
step: 80, loss: 0.3145616948604584
step: 90, loss: 0.2263065129518509
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.18601161241531372
step: 110, loss: 0.45919299125671387
step: 120, loss: 0.3072752356529236
step: 130, loss: 0.30391526222229004
step: 140, loss: 0.40033474564552307
step: 150, loss: 0.32915881276130676
step: 160, loss: 0.39398711919784546
step: 170, loss: 0.37990936636924744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.22583025830258302, f1=0.23288749016522423,
best_f1=0.23288749016522423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30041882395744324
step: 10, loss: 0.4603578746318817
step: 20, loss: 0.3265819847583771
step: 30, loss: 0.2823317050933838
step: 40, loss: 0.06448150426149368
step: 50, loss: 0.38751375675201416
step: 60, loss: 0.19500736892223358
step: 70, loss: 0.5188543796539307
step: 80, loss: 0.2362861931324005
step: 90, loss: 0.21150387823581696
step: 100, loss: 0.4813205599784851
step: 110, loss: 0.2717757225036621
step: 120, loss: 0.2472497522830963
step: 130, loss: 0.4917217791080475
step: 140, loss: 0.4566327631473541
step: 150, loss: 0.38037240505218506
step: 160, loss: 0.3895992934703827
step: 170, loss: 0.34467387199401855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3437037037037037, f1=0.2940275650842266,
best_f1=0.2940275650842266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5052491426467896
step: 10, loss: 0.2768993377685547
step: 20, loss: 0.17663919925689697
step: 30, loss: 0.16309981048107147
step: 40, loss: 0.22589151561260223
step: 50, loss: 0.3154698312282562
step: 60, loss: 0.09683256596326828
step: 70, loss: 0.08606822788715363
step: 80, loss: 0.11604221165180206
step: 90, loss: 0.29850372672080994
step: 100, loss: 0.15392285585403442
step: 110, loss: 0.034586865454912186
step: 120, loss: 0.25226491689682007

```

```

step: 130, loss: 0.3416345715522766
step: 140, loss: 0.19810910522937775
step: 150, loss: 0.09070739895105362
step: 160, loss: 0.17578111588954926
step: 170, loss: 0.08090205490589142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.762135922330097, f1=0.7431192660550459,
best_f1=0.7431192660550459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12047816812992096
step: 10, loss: 0.26563581824302673
step: 20, loss: 0.07371113449335098
step: 30, loss: 0.35705500841140747
step: 40, loss: 0.11938071995973587
step: 50, loss: 0.10071214288473129
step: 60, loss: 0.12714359164237976
step: 70, loss: 0.011842818930745125
step: 80, loss: 0.20913608372211456
step: 90, loss: 0.24236029386520386
step: 100, loss: 0.22656752169132233
step: 110, loss: 0.21965491771697998
step: 120, loss: 0.3510839343070984
step: 130, loss: 0.1088109016418457
step: 140, loss: 0.09774449467658997
step: 150, loss: 0.18049485981464386
step: 160, loss: 0.006339434068650007
step: 170, loss: 0.09733190387487411
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8010899182561307, f1=0.8167539267015707,
best_f1=0.8167539267015707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01301082968711853

```

```

step: 10, loss: 0.05460033938288689
step: 20, loss: 0.15217113494873047
step: 30, loss: 0.05836839973926544
step: 40, loss: 0.12562404572963715
step: 50, loss: 0.03792903199791908
step: 60, loss: 0.04975986108183861
step: 70, loss: 0.10657431930303574
step: 80, loss: 0.04616639018058777
step: 90, loss: 0.12550747394561768
step: 100, loss: 0.13342179358005524
step: 110, loss: 0.16093234717845917
step: 120, loss: 0.12225498259067535
step: 130, loss: 0.006952665280550718
step: 140, loss: 0.06759874522686005
step: 150, loss: 0.15128983557224274
step: 160, loss: 0.04399101063609123
step: 170, loss: 0.016158636659383774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8059701492537313, f1=0.827906976744186,
best_f1=0.827906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15959347784519196
step: 10, loss: 0.07694220542907715
step: 20, loss: 0.09429577738046646
step: 30, loss: 0.07446962594985962
step: 40, loss: 0.05789335444569588
step: 50, loss: 0.03775367885828018
step: 60, loss: 0.024861304089426994
step: 70, loss: 0.03891767933964729
step: 80, loss: 0.02393394522368908
step: 90, loss: 0.2097969502210617
step: 100, loss: 0.016705557703971863
step: 110, loss: 0.2147648185491562
step: 120, loss: 0.06702828407287598
step: 130, loss: 0.03949548676609993
step: 140, loss: 0.21379640698432922
step: 150, loss: 0.023378584533929825
step: 160, loss: 0.054325979202985764
step: 170, loss: 0.04227786883711815
/usr/local/lib/python3.7/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=0.8021390374331551, f1=0.8549618320610687,
best_f1=0.827906976744186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025225814431905746
step: 10, loss: 0.01407978031784296
step: 20, loss: 0.038736023008823395
step: 30, loss: 0.0042547970078885555
step: 40, loss: 0.0017542734276503325
step: 50, loss: 0.002612409181892872
step: 60, loss: 0.08327414095401764
step: 70, loss: 0.004219340160489082
step: 80, loss: 0.02454276755452156
step: 90, loss: 0.11525610834360123
step: 100, loss: 0.00516539579257369
step: 110, loss: 0.11799495667219162
step: 120, loss: 0.019351208582520485
step: 130, loss: 0.039341703057289124
step: 140, loss: 0.005128142423927784
step: 150, loss: 0.16033770143985748
step: 160, loss: 0.01605338230729103
step: 170, loss: 0.0291235763579607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8186528497409326, f1=0.8543689320388349,
best_f1=0.8543689320388349
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0734369084239006
step: 10, loss: 0.11446475982666016
step: 20, loss: 0.008213434368371964
step: 30, loss: 0.033643778413534164
step: 40, loss: 0.005846993997693062
step: 50, loss: 0.0027933320961892605
step: 60, loss: 0.0114052202552557
step: 70, loss: 0.055368296802043915
step: 80, loss: 0.15344823896884918
step: 90, loss: 0.021097516641020775
step: 100, loss: 0.003747184993699193

```



```

step: 110, loss: 0.04148448631167412
step: 120, loss: 0.06727329641580582
step: 130, loss: 0.011813986115157604
step: 140, loss: 0.0843772366642952
step: 150, loss: 0.003399674082174897
step: 160, loss: 0.0087616927921772
step: 170, loss: 0.13864338397979736
/usr/local/lib/python3.7/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=0.8031088082901555, f1=0.8341463414634146,
best_f1=0.8543689320388349
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11607745289802551
step: 10, loss: 0.010819879360496998
step: 20, loss: 0.011445505544543266
step: 30, loss: 0.004137223120778799
step: 40, loss: 0.12252148985862732
step: 50, loss: 0.05538045987486839
step: 60, loss: 0.0064909146167337894
step: 70, loss: 0.23474299907684326
step: 80, loss: 0.001228830311447382
step: 90, loss: 0.007221736013889313
step: 100, loss: 0.025072623044252396
step: 110, loss: 0.03539882227778435
step: 120, loss: 0.08996757864952087
step: 130, loss: 0.0032665501348674297
step: 140, loss: 0.015521487221121788
step: 150, loss: 0.23343104124069214
step: 160, loss: 0.009383008815348148
step: 170, loss: 0.0504855141043663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8282828282828283, f1=0.8611764705882352,
best_f1=0.8611764705882352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026935571804642677
step: 10, loss: 0.005372363142669201
step: 20, loss: 0.15619730949401855

```

```

step: 30, loss: 0.0768110603094101
step: 40, loss: 0.012132818810641766
step: 50, loss: 0.008039248175919056
step: 60, loss: 0.10892630368471146
step: 70, loss: 0.008572684600949287
step: 80, loss: 0.01306848507374525
step: 90, loss: 0.009064484387636185
step: 100, loss: 0.004086955450475216
step: 110, loss: 0.005993728060275316
step: 120, loss: 0.0009151023114100099
step: 130, loss: 0.10212939977645874
step: 140, loss: 0.02800852619111538
step: 150, loss: 0.011873620562255383
step: 160, loss: 0.001414285390637815
step: 170, loss: 0.05253363028168678
/usr/local/lib/python3.7/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=0.8229166666666666, f1=0.8585365853658536,
best_f1=0.8611764705882352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05170361325144768
step: 10, loss: 0.00038227386539801955
step: 20, loss: 0.0018336961511522532
step: 30, loss: 0.05344843491911888
step: 40, loss: 0.0012808690080419183
step: 50, loss: 0.029004013165831566
step: 60, loss: 0.0014374819584190845
step: 70, loss: 0.006809718441218138
step: 80, loss: 0.10010501742362976
step: 90, loss: 0.009056570008397102
step: 100, loss: 0.09657931327819824
step: 110, loss: 0.0013067851541563869
step: 120, loss: 0.002107941312715411
step: 130, loss: 0.01711299642920494
step: 140, loss: 0.005149866919964552
step: 150, loss: 0.00039045122684910893
step: 160, loss: 0.09511922299861908
step: 170, loss: 0.03427819535136223
/usr/local/lib/python3.7/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=0.8169761273209549, f1=0.8746928746928747,
best_f1=0.8611764705882352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001899448106996715
step: 10, loss: 0.001283489167690277
step: 20, loss: 0.008740625344216824
step: 30, loss: 0.08378716558218002
step: 40, loss: 0.005793221294879913
step: 50, loss: 0.0006360645056702197
step: 60, loss: 0.012533271685242653
step: 70, loss: 0.0004963345709256828
step: 80, loss: 0.0006516216089949012
step: 90, loss: 0.02183631621301174
step: 100, loss: 0.004088564310222864
step: 110, loss: 0.06945479661226273
step: 120, loss: 0.014567849226295948
step: 130, loss: 0.03102191723883152
step: 140, loss: 0.000775067659560591
step: 150, loss: 0.0022571750450879335
step: 160, loss: 0.0987800732254982
step: 170, loss: 0.0004798536829184741
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8307692307692308, f1=0.8598574821852732,
best_f1=0.8598574821852732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002620006212964654
step: 10, loss: 0.0004204793367534876
step: 20, loss: 0.0009776740334928036
step: 30, loss: 0.022523611783981323
step: 40, loss: 0.0005726031376980245
step: 50, loss: 0.005871837493032217
step: 60, loss: 0.0005700662732124329
step: 70, loss: 0.06307699531316757
step: 80, loss: 0.0005721044726669788
step: 90, loss: 0.001302272779867053
step: 100, loss: 0.04955461993813515
step: 110, loss: 0.007603551726788282
step: 120, loss: 0.0045636966824531555
step: 130, loss: 0.00041789939859881997
step: 140, loss: 0.0034442690666764975
step: 150, loss: 0.0003986287338193506
step: 160, loss: 0.0038044736720621586

```

```

step: 170, loss: 0.00020496503566391766
/usr/local/lib/python3.7/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=0.8116710875331565, f1=0.8522167487684729,
best_f1=0.8598574821852732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03435422107577324
step: 10, loss: 0.000392824673326686
step: 20, loss: 0.017775055021047592
step: 30, loss: 0.0005326417740434408
step: 40, loss: 0.0003083463234361261
step: 50, loss: 0.0025803386233747005
step: 60, loss: 0.0020161569118499756
step: 70, loss: 0.006064675282686949
step: 80, loss: 0.0013242976274341345
step: 90, loss: 0.0004016056191176176
step: 100, loss: 0.005165076814591885
step: 110, loss: 0.031462207436561584
step: 120, loss: 0.0018212939612567425
step: 130, loss: 0.022490451112389565
step: 140, loss: 0.018810739740729332
step: 150, loss: 0.019428126513957977
step: 160, loss: 0.0004382760380394757
step: 170, loss: 0.021483486518263817
/usr/local/lib/python3.7/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=0.8160000000000001, f1=0.8734491315136476,
best_f1=0.8598574821852732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002542816800996661
step: 10, loss: 0.004852264188230038
step: 20, loss: 0.0005219242884777486
step: 30, loss: 0.0013174251653254032
step: 40, loss: 0.004309887997806072
step: 50, loss: 0.0014648468932136893
step: 60, loss: 0.002779686823487282
step: 70, loss: 0.0924164429306984
step: 80, loss: 0.0010009582620114088
step: 90, loss: 0.0013410894898697734
step: 100, loss: 0.02852698601782322
step: 110, loss: 0.051654152572155
step: 120, loss: 0.0005404429393820465

```

```

step: 130, loss: 0.002090127905830741
step: 140, loss: 0.0015838340623304248
step: 150, loss: 0.003891501110047102
step: 160, loss: 0.0009942465694621205
step: 170, loss: 0.0005337355541996658
/usr/local/lib/python3.7/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=0.8157248157248157, f1=0.8387096774193549,
best_f1=0.8598574821852732
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, 317.37it/s]
load_f1 = 0.823529411764706
real_f1 = 0.820253164556962
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 196.24it/s]

```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5877136588096619
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.5096930265426636
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: 20, loss: 0.494449645280838
step: 30, loss: 0.37452414631843567
step: 40, loss: 0.3456761837005615
step: 50, loss: 0.5775570273399353
step: 60, loss: 0.3281842768192291
step: 70, loss: 0.09120471030473709
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 80, loss: 0.3984716832637787
step: 90, loss: 0.33508485555648804
step: 100, loss: 0.12574680149555206
step: 110, loss: 0.21259772777557373
step: 120, loss: 0.09679455310106277
step: 130, loss: 0.06765156239271164
step: 140, loss: 0.015546015463769436
step: 150, loss: 0.1307346224784851
step: 160, loss: 0.012618055567145348
step: 170, loss: 0.1264895498752594
step: 180, loss: 0.1451507955789566
step: 190, loss: 0.09769698977470398
step: 200, loss: 0.027505742385983467
step: 210, loss: 0.04476471617817879
step: 220, loss: 0.0961838811635971
step: 230, loss: 0.014672130346298218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9576837416481068, f1=0.9571106094808126,
best_f1=0.9571106094808126

```

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

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

```
step: 0, loss: 0.023018090054392815  
step: 10, loss: 0.16577263176441193  
step: 20, loss: 0.03303293138742447  
step: 30, loss: 0.04094882681965828  
step: 40, loss: 0.07539165019989014  
step: 50, loss: 0.033481817692518234  
step: 60, loss: 0.014128969050943851  
step: 70, loss: 0.039335936307907104  
step: 80, loss: 0.0034888789523392916  
step: 90, loss: 0.015921849757432938  
step: 100, loss: 0.00883724819868803  
step: 110, loss: 0.0169843677431345  
step: 120, loss: 0.052062951028347015  
step: 130, loss: 0.06941819190979004  
step: 140, loss: 0.015131951309740543  
step: 150, loss: 0.10249930620193481  
step: 160, loss: 0.0046950713731348515  
step: 170, loss: 0.001300669857300818  
step: 180, loss: 0.1206674724817276  
step: 190, loss: 0.08625654876232147  
step: 200, loss: 0.03397028148174286  
step: 210, loss: 0.05724157392978668  
step: 220, loss: 0.006403648294508457  
step: 230, loss: 0.0020589795894920826
```

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

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

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

```
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=0.9655172413793103, f1=0.967525195968645,  
best_f1=0.967525195968645
```

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

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

```
step: 0, loss: 0.028005125001072884  
step: 10, loss: 0.0024259702768176794  
step: 20, loss: 0.07691022753715515  
step: 30, loss: 0.003351405495777726  
step: 40, loss: 0.002811113139614463  
step: 50, loss: 0.02397708222270012  
step: 60, loss: 0.02323134057223797  
step: 70, loss: 0.010218090377748013  
step: 80, loss: 0.09569089114665985
```



```

step: 90, loss: 0.0012515031266957521
step: 100, loss: 0.012495146133005619
step: 110, loss: 0.0016943068476393819
step: 120, loss: 0.0009050755179487169
step: 130, loss: 0.08690774440765381
step: 140, loss: 0.004636828787624836
step: 150, loss: 0.022619495168328285
step: 160, loss: 0.003424334339797497
step: 170, loss: 0.000649425492156297
step: 180, loss: 0.005718957167118788
step: 190, loss: 0.21980920433998108
step: 200, loss: 0.013788867741823196
step: 210, loss: 0.00476105697453022
step: 220, loss: 0.1004248782992363
step: 230, loss: 0.052487313747406006
/usr/local/lib/python3.7/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=0.9564245810055866, f1=0.9588431590656283,
best_f1=0.967525195968645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019833935424685478
step: 10, loss: 0.0013759235152974725
step: 20, loss: 0.013724585995078087
step: 30, loss: 0.09179113060235977
step: 40, loss: 0.06748169660568237
step: 50, loss: 0.0394999124109745
step: 60, loss: 0.023395061492919922
step: 70, loss: 0.0736980065703392
step: 80, loss: 0.10186921805143356
step: 90, loss: 0.004532447550445795
step: 100, loss: 0.0549132414162159
step: 110, loss: 0.0010853934800252318
step: 120, loss: 0.006218448281288147
step: 130, loss: 0.003575833048671484
step: 140, loss: 0.000894277123734355
step: 150, loss: 0.007869108580052853
step: 160, loss: 0.0015699077630415559
step: 170, loss: 0.001508429297246039
step: 180, loss: 0.10415898263454437
step: 190, loss: 0.0031355747487396
step: 200, loss: 0.012914341874420643
step: 210, loss: 0.013233947567641735
step: 220, loss: 0.0008018460939638317
step: 230, loss: 0.004494486842304468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9722530521642618, f1=0.968609865470852,
best_f1=0.968609865470852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009024813771247864
step: 10, loss: 0.0011674639536067843
step: 20, loss: 0.08917966485023499
step: 30, loss: 0.000514661893248558
step: 40, loss: 0.005835245829075575
step: 50, loss: 0.0012675668112933636
step: 60, loss: 0.003845298197120428
step: 70, loss: 0.00040511388215236366
step: 80, loss: 0.09086204320192337
step: 90, loss: 0.02446696162223816
step: 100, loss: 0.00031890146783553064
step: 110, loss: 0.0006134259747341275
step: 120, loss: 0.021596498787403107
step: 130, loss: 0.007783340290188789
step: 140, loss: 0.003914008382707834
step: 150, loss: 0.009876578114926815
step: 160, loss: 0.00034982486977241933
step: 170, loss: 0.035079117864370346
step: 180, loss: 0.018394194543361664
step: 190, loss: 0.01922653056681156
step: 200, loss: 0.006481243297457695
step: 210, loss: 0.0015844710869714618
step: 220, loss: 0.002312404802069068
step: 230, loss: 0.0330367237329483
/usr/local/lib/python3.7/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=0.9685393258426966, f1=0.975, best_f1=0.968609865470852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005942843854427338
step: 10, loss: 0.003342239186167717
step: 20, loss: 0.0025191777385771275
step: 30, loss: 0.0008096863166429102
step: 40, loss: 0.0005159182474017143
step: 50, loss: 0.0005265157669782639

```

```

step: 60, loss: 0.0014516719384118915
step: 70, loss: 0.13787657022476196
step: 80, loss: 0.002164439996704459
step: 90, loss: 0.05657775700092316
step: 100, loss: 0.0013674385845661163
step: 110, loss: 0.006027556024491787
step: 120, loss: 0.007499438710510731
step: 130, loss: 0.002328475471585989
step: 140, loss: 0.0005299651529639959
step: 150, loss: 0.00021798866509925574
step: 160, loss: 0.013480191119015217
step: 170, loss: 0.0008196835406124592
step: 180, loss: 0.00024417528766207397
step: 190, loss: 0.0005578339332714677
step: 200, loss: 0.0008021878893487155
step: 210, loss: 0.001162012224085629
step: 220, loss: 0.00035244025639258325
step: 230, loss: 0.0015100965974852443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9752252252252253, f1=0.9739524348810873,
best_f1=0.9739524348810873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004446233797352761
step: 10, loss: 0.0002727157552726567
step: 20, loss: 0.026834890246391296
step: 30, loss: 0.0030160725582391024
step: 40, loss: 0.0012656105682253838
step: 50, loss: 0.00312465475872159
step: 60, loss: 0.0006491553504019976
step: 70, loss: 0.0003977838787250221
step: 80, loss: 0.023055074736475945
step: 90, loss: 0.12602774798870087
step: 100, loss: 0.02624642290174961
step: 110, loss: 0.00046330876648426056
step: 120, loss: 0.0004607374139595777
step: 130, loss: 0.000345437292708084
step: 140, loss: 0.00018012650252785534
step: 150, loss: 0.0034481242764741182
step: 160, loss: 0.0005895631620660424
step: 170, loss: 0.0006542180781252682

```

```

step: 180, loss: 0.0001812139671528712
step: 190, loss: 0.021496227011084557
step: 200, loss: 0.0016418768791481853
step: 210, loss: 0.0009064944460988045
step: 220, loss: 0.0005260384059511125
step: 230, loss: 0.00040870060911402106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9764837625979844, f1=0.9796839729119639,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035343310446478426
step: 10, loss: 0.005966638680547476
step: 20, loss: 0.00028479244792833924
step: 30, loss: 0.0005281370831653476
step: 40, loss: 0.0009079111041501164
step: 50, loss: 0.00019544563838280737
step: 60, loss: 0.0002526319003663957
step: 70, loss: 0.00024305927217938006
step: 80, loss: 0.024593116715550423
step: 90, loss: 0.0007204597350209951
step: 100, loss: 0.0004987421561963856
step: 110, loss: 0.0002687597298063338
step: 120, loss: 0.00040790982893668115
step: 130, loss: 0.000873288547154516
step: 140, loss: 0.00017929122259374708
step: 150, loss: 0.04129507765173912
step: 160, loss: 0.00018475332763046026
step: 170, loss: 0.003810272319242358
step: 180, loss: 0.0005309790722094476
step: 190, loss: 0.0014106814051046968
step: 200, loss: 0.002982160309329629
step: 210, loss: 0.0009337192750535905
step: 220, loss: 0.0005426608840934932
step: 230, loss: 0.0011062206467613578
/usr/local/lib/python3.7/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=0.9741863075196409, f1=0.9772727272727272,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031524328514933586
step: 10, loss: 0.00048160713049583137
step: 20, loss: 0.027458228170871735
step: 30, loss: 0.0016776026459410787
step: 40, loss: 0.0003581398632377386
step: 50, loss: 0.006678285077214241
step: 60, loss: 0.0006573303253389895
step: 70, loss: 0.035160671919584274
step: 80, loss: 0.00019954283197876066
step: 90, loss: 0.019565308466553688
step: 100, loss: 0.0002613025135360658
step: 110, loss: 0.00015375415387097746
step: 120, loss: 0.030326060950756073
step: 130, loss: 0.000539321918040514
step: 140, loss: 0.0003401818103156984
step: 150, loss: 0.0027950124349445105
step: 160, loss: 0.0009486464550718665
step: 170, loss: 0.00016909754776861519
step: 180, loss: 0.0005385483964346349
step: 190, loss: 0.00015624964726157486
step: 200, loss: 0.011028178967535496
step: 210, loss: 0.006746073253452778
step: 220, loss: 0.0004831335972994566
step: 230, loss: 0.08717446774244308
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9774774774774775, f1=0.971815107102593,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016998662613332272
step: 10, loss: 0.0003463400644250214
step: 20, loss: 0.00015951313253026456
step: 30, loss: 0.0001480564969824627
step: 40, loss: 0.005327858030796051
step: 50, loss: 7.94864390627481e-05
step: 60, loss: 0.0003209470596630126
step: 70, loss: 0.009394421242177486
step: 80, loss: 0.0002458761155139655
step: 90, loss: 0.00022465302026830614

```

```

step: 100, loss: 0.004388428293168545
step: 110, loss: 0.0001715072721708566
step: 120, loss: 0.00015522888861596584
step: 130, loss: 0.00024099150323309004
step: 140, loss: 0.0003010024956893176
step: 150, loss: 0.00031105964444577694
step: 160, loss: 7.619996904395521e-05
step: 170, loss: 0.0001522455713711679
step: 180, loss: 0.012964536435902119
step: 190, loss: 0.00020988541655242443
step: 200, loss: 0.00028802824090234935
step: 210, loss: 0.012217643670737743
step: 220, loss: 0.00023370562121272087
step: 230, loss: 6.371392373694107e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9810901001112348, f1=0.9733333333333333,
best_f1=0.9733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.955886445008218e-05
step: 10, loss: 0.00012178994074929506
step: 20, loss: 0.00013762034359388053
step: 30, loss: 0.00039448851021006703
step: 40, loss: 3.6471421481110156e-05
step: 50, loss: 7.029239350231364e-05
step: 60, loss: 0.0005065381992608309
step: 70, loss: 9.36813376029022e-05
step: 80, loss: 0.008386398665606976
step: 90, loss: 0.0005790256545878947
step: 100, loss: 0.0002122296573361382
step: 110, loss: 0.00019290570344310254
step: 120, loss: 0.0002803737297654152
step: 130, loss: 0.00015541173343081027
step: 140, loss: 0.0001863515644799918
step: 150, loss: 0.00020090583711862564
step: 160, loss: 0.014931539073586464
step: 170, loss: 0.00038909740396775305
step: 180, loss: 0.0002687326923478395
step: 190, loss: 0.000499385001603514
step: 200, loss: 0.0006289728335104883
step: 210, loss: 0.00019600876839831471

```

```

step: 220, loss: 0.00021003664005547762
step: 230, loss: 0.00010122806997969747
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.978675645342312, f1=0.9808342728297633,
best_f1=0.9733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013756711268797517
step: 10, loss: 7.333321264013648e-05
step: 20, loss: 0.009568026289343834
step: 30, loss: 0.003516382072120905
step: 40, loss: 0.00012663326924666762
step: 50, loss: 0.00012945369235239923
step: 60, loss: 0.00011589685163926333
step: 70, loss: 9.483149187872186e-05
step: 80, loss: 7.642410491826013e-05
step: 90, loss: 9.533196862321347e-05
step: 100, loss: 7.117628410924226e-05
step: 110, loss: 7.260978600243106e-05
step: 120, loss: 0.0005850230227224529
step: 130, loss: 8.453439659206197e-05
step: 140, loss: 8.617062121629715e-05
step: 150, loss: 7.485061360057443e-05
step: 160, loss: 0.00047427439130842686
step: 170, loss: 9.950107050826773e-05
step: 180, loss: 5.981310096103698e-05
step: 190, loss: 0.00015018567501101643
step: 200, loss: 7.069831917760894e-05
step: 210, loss: 6.242904055397958e-05
step: 220, loss: 0.011156356893479824
step: 230, loss: 7.882279169280082e-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=0.9808773903262092, f1=0.9773242630385486,
best_f1=0.9733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.747632480459288e-05
step: 10, loss: 7.727531192358583e-05
step: 20, loss: 8.999454439617693e-05
step: 30, loss: 3.87860054615885e-05
step: 40, loss: 8.717804303159937e-05
step: 50, loss: 0.001804278464987874

```

```

step: 60, loss: 5.944490840192884e-05
step: 70, loss: 5.798863639938645e-05
step: 80, loss: 3.41027116519399e-05
step: 90, loss: 4.9583188229007646e-05
step: 100, loss: 7.524860120611265e-05
step: 110, loss: 8.137567056110129e-05
step: 120, loss: 8.316911407746375e-05
step: 130, loss: 6.479456351371482e-05
step: 140, loss: 0.0007594986818730831
step: 150, loss: 5.468802191899158e-05
step: 160, loss: 0.0004862291971221566
step: 170, loss: 8.006584539543837e-05
step: 180, loss: 0.016924714669585228
step: 190, loss: 6.86286948621273e-05
step: 200, loss: 5.033369598095305e-05
step: 210, loss: 0.0002599881845526397
step: 220, loss: 6.591275450773537e-05
step: 230, loss: 8.597939449828118e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9819819819819819, f1=0.9773242630385486,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.1533843614161015e-05
step: 10, loss: 6.863485759822652e-05
step: 20, loss: 5.860820965608582e-05
step: 30, loss: 7.599150558235124e-05
step: 40, loss: 8.191597589757293e-05
step: 50, loss: 5.0702932639978826e-05
step: 60, loss: 6.815343658672646e-05
step: 70, loss: 0.00020448144641704857
step: 80, loss: 4.883009023615159e-05
step: 90, loss: 7.339727744692937e-05
step: 100, loss: 5.53043901163619e-05
step: 110, loss: 6.609151023440063e-05
step: 120, loss: 3.773695425479673e-05
step: 130, loss: 0.0001255266834050417
step: 140, loss: 6.221909279702231e-05
step: 150, loss: 7.060926873236895e-05
step: 160, loss: 6.30128342891112e-05
step: 170, loss: 0.0029833484441041946

```



```

step: 180, loss: 4.767270365846343e-05
step: 190, loss: 4.465167876332998e-05
step: 200, loss: 7.798853766871616e-05
step: 210, loss: 4.775397974299267e-05
step: 220, loss: 6.306942668743432e-05
step: 230, loss: 2.816498999891337e-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=0.9819819819819819, f1=0.976271186440678,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011636155977612361
step: 10, loss: 0.0001587720907991752
step: 20, loss: 7.464185182470828e-05
step: 30, loss: 6.213428423507139e-05
step: 40, loss: 3.568228930816986e-05
step: 50, loss: 5.551180220209062e-05
step: 60, loss: 0.05275333672761917
step: 70, loss: 4.603292836691253e-05
step: 80, loss: 5.642844917019829e-05
step: 90, loss: 4.963060928275809e-05
step: 100, loss: 3.380857378942892e-05
step: 110, loss: 5.727943789679557e-05
step: 120, loss: 0.00019278854597359896
step: 130, loss: 4.5910506742075086e-05
step: 140, loss: 0.005656686145812273
step: 150, loss: 0.0016967010451480746
step: 160, loss: 5.000276360078715e-05
step: 170, loss: 2.402694917691406e-05
step: 180, loss: 5.815381882712245e-05
step: 190, loss: 3.578920222935267e-05
step: 200, loss: 0.0026441419031471014
step: 210, loss: 0.0031571644358336926
step: 220, loss: 5.009715096093714e-05
step: 230, loss: 6.463795580202714e-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=0.9808773903262092, f1=0.9786276715410572,
best_f1=0.9773242630385486
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 205.94it/s]
load_f1 = 0.9808773903262092
real_f1 = 0.9808342728297633
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 195.35it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--lm roberta \
--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.631782591342926

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

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

step: 10, loss: 0.4555045962333679

step: 20, loss: 0.333138644695282
step: 30, loss: 0.37839755415916443
step: 40, loss: 0.3488781452178955
step: 50, loss: 0.5098021626472473
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.19308900833129883
step: 70, loss: 0.44346097111701965
step: 80, loss: 0.3446418344974518
step: 90, loss: 0.26860520243644714
step: 100, loss: 0.4923666715621948
step: 110, loss: 0.2184096872806549
step: 120, loss: 0.36984333395957947
step: 130, loss: 0.15548454225063324
step: 140, loss: 0.2006170153617859
step: 150, loss: 0.10612406581640244
step: 160, loss: 0.22703132033348083
step: 170, loss: 0.17014563083648682
step: 180, loss: 0.0702565535902977
step: 190, loss: 0.10846435278654099
step: 200, loss: 0.09909257292747498
step: 210, loss: 0.03780537098646164
step: 220, loss: 0.1569242775440216
step: 230, loss: 0.07396862655878067
step: 240, loss: 0.03900916501879692
step: 250, loss: 0.0580935962498188
step: 260, loss: 0.3119152784347534
step: 270, loss: 0.3404353857040405
step: 280, loss: 0.09591425210237503
step: 290, loss: 0.10609687119722366
step: 300, loss: 0.3618302643299103
step: 310, loss: 0.14497707784175873
step: 320, loss: 0.07544904202222824
step: 330, loss: 0.10870999842882156
step: 340, loss: 0.38855358958244324
step: 350, loss: 0.2078431248664856
step: 360, loss: 0.06214771419763565
step: 370, loss: 0.04404650628566742
step: 380, loss: 0.1706133335828781
step: 390, loss: 0.07453544437885284
step: 400, loss: 0.06455038487911224
step: 410, loss: 0.25461477041244507
step: 420, loss: 0.02193339914083481
step: 430, loss: 0.022801784798502922
step: 440, loss: 0.0728413388133049
step: 450, loss: 0.157437264919281
step: 460, loss: 0.06376373767852783
step: 470, loss: 0.14149723947048187
step: 480, loss: 0.12116976082324982

```

step: 490, loss: 0.12804877758026123
step: 500, loss: 0.04884662851691246
step: 510, loss: 0.09752456098794937
step: 520, loss: 0.25047746300697327
step: 530, loss: 0.09042761474847794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9175694771549693, f1=0.9260808926080892,
best_f1=0.9260808926080892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11410371959209442
step: 10, loss: 0.1320858895778656
step: 20, loss: 0.1374908834695816
step: 30, loss: 0.29831427335739136
step: 40, loss: 0.15608201920986176
step: 50, loss: 0.10340280830860138
step: 60, loss: 0.012797173112630844
step: 70, loss: 0.046139806509017944
step: 80, loss: 0.19974954426288605
step: 90, loss: 0.034690581262111664
step: 100, loss: 0.16520090401172638
step: 110, loss: 0.02541961334645748
step: 120, loss: 0.08511444181203842
step: 130, loss: 0.011751316487789154
step: 140, loss: 0.031377557665109634
step: 150, loss: 0.03185281157493591
step: 160, loss: 0.08435797691345215
step: 170, loss: 0.03556640073657036
step: 180, loss: 0.16456329822540283
step: 190, loss: 0.01602240838110447
step: 200, loss: 0.18560895323753357
step: 210, loss: 0.015890054404735565
step: 220, loss: 0.0014647983480244875
step: 230, loss: 0.0985281839966774
step: 240, loss: 0.017192451283335686
step: 250, loss: 0.1038811132311821
step: 260, loss: 0.06547028571367264
step: 270, loss: 0.07108259946107864
step: 280, loss: 0.0156598761677742
step: 290, loss: 0.03526689484715462
step: 300, loss: 0.09760074317455292

```

```

step: 310, loss: 0.09802591055631638
step: 320, loss: 0.07924991846084595
step: 330, loss: 0.14608162641525269
step: 340, loss: 0.13134710490703583
step: 350, loss: 0.0029443365056067705
step: 360, loss: 0.020737916231155396
step: 370, loss: 0.015570506453514099
step: 380, loss: 0.18435725569725037
step: 390, loss: 0.013412238098680973
step: 400, loss: 0.0954323559999466
step: 410, loss: 0.019988883286714554
step: 420, loss: 0.2891899347305298
step: 430, loss: 0.05089844390749931
step: 440, loss: 0.003809294430539012
step: 450, loss: 0.018127264454960823
step: 460, loss: 0.04023610055446625
step: 470, loss: 0.09454599022865295
step: 480, loss: 0.011327706277370453
step: 490, loss: 0.12164454162120819
step: 500, loss: 0.006209196988493204
step: 510, loss: 0.024924172088503838
step: 520, loss: 0.31810668110847473
step: 530, loss: 0.07969043403863907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.939435968562182, f1=0.9282385834109972,
best_f1=0.9282385834109972
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13334837555885315
step: 10, loss: 0.0684908851981163
step: 20, loss: 0.006859467830508947
step: 30, loss: 0.057125844061374664
step: 40, loss: 0.057654593139886856
step: 50, loss: 0.01144950557500124
step: 60, loss: 0.04918099567294121
step: 70, loss: 0.020785709843039513
step: 80, loss: 0.03961103782057762
step: 90, loss: 0.008393206633627415
step: 100, loss: 0.010241793468594551
step: 110, loss: 0.004577663727104664
step: 120, loss: 0.19488845765590668

```

```

step: 130, loss: 0.0755421370267868
step: 140, loss: 0.043139226734638214
step: 150, loss: 0.025840651243925095
step: 160, loss: 0.011778856627643108
step: 170, loss: 0.0033826034050434828
step: 180, loss: 0.04834707826375961
step: 190, loss: 0.00590540561825037
step: 200, loss: 0.027004053816199303
step: 210, loss: 0.06572563201189041
step: 220, loss: 0.13376782834529877
step: 230, loss: 0.07570237666368484
step: 240, loss: 0.04966733604669571
step: 250, loss: 0.06796569377183914
step: 260, loss: 0.1329449862241745
step: 270, loss: 0.007863745093345642
step: 280, loss: 0.012094888836145401
step: 290, loss: 0.024912578985095024
step: 300, loss: 0.08657203614711761
step: 310, loss: 0.1325453370809555
step: 320, loss: 0.09236380457878113
step: 330, loss: 0.008914279751479626
step: 340, loss: 0.00925588607788086
step: 350, loss: 0.2214168757200241
step: 360, loss: 0.017191695049405098
step: 370, loss: 0.04220151901245117
step: 380, loss: 0.00903654657304287
step: 390, loss: 0.022901132702827454
step: 400, loss: 0.19350001215934753
step: 410, loss: 0.04551893100142479
step: 420, loss: 0.027932794764637947
step: 430, loss: 0.03968053683638573
step: 440, loss: 0.1293097734451294
step: 450, loss: 0.051552701741456985
step: 460, loss: 0.07401905953884125
step: 470, loss: 0.05446178466081619
step: 480, loss: 0.32243451476097107
step: 490, loss: 0.02722114510834217
step: 500, loss: 0.03208146244287491
step: 510, loss: 0.026121899485588074
step: 520, loss: 0.0011206648778170347
step: 530, loss: 0.004375522956252098
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9397146801656695, f1=0.9328426862925483,  
best_f1=0.9328426862925483  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.12895238399505615  
step: 10, loss: 0.006196311675012112  
step: 20, loss: 0.0328909195959568  
step: 30, loss: 0.08953876793384552  
step: 40, loss: 0.01113420445472002  
step: 50, loss: 0.008293329738080502  
step: 60, loss: 0.007267045322805643  
step: 70, loss: 0.005694454535841942  
step: 80, loss: 0.12202037125825882  
step: 90, loss: 0.017084116116166115  
step: 100, loss: 0.0017700850730761886  
step: 110, loss: 0.10707930475473404  
step: 120, loss: 0.02510775253176689  
step: 130, loss: 0.18749214708805084  
step: 140, loss: 0.021308720111846924  
step: 150, loss: 0.011452102102339268  
step: 160, loss: 0.005486357491463423  
step: 170, loss: 0.010123453103005886  
step: 180, loss: 0.04070258513092995  
step: 190, loss: 0.048374734818935394  
step: 200, loss: 0.038033582270145416  
step: 210, loss: 0.002113067079335451  
step: 220, loss: 0.016839880496263504  
step: 230, loss: 0.0020324736833572388  
step: 240, loss: 0.002618519589304924  
step: 250, loss: 0.03085421398282051  
step: 260, loss: 0.016485117375850677  
step: 270, loss: 0.12065991759300232  
step: 280, loss: 0.004795708227902651  
step: 290, loss: 0.020063983276486397  
step: 300, loss: 0.004173988476395607  
step: 310, loss: 0.005864561535418034  
step: 320, loss: 0.15200607478618622  
step: 330, loss: 0.007813737727701664  
step: 340, loss: 0.005564059130847454  
step: 350, loss: 0.08339434117078781  
step: 360, loss: 0.07311940938234329  
step: 370, loss: 0.003686177311465144  
step: 380, loss: 0.002391423797234893  
step: 390, loss: 0.0010984714608639479  
step: 400, loss: 0.007403066847473383  
step: 410, loss: 0.0009518942097201943  
step: 420, loss: 0.024741029366850853
```



```

step: 430, loss: 0.003604459809139371
step: 440, loss: 0.003926175180822611
step: 450, loss: 0.01853948086500168
step: 460, loss: 0.1594555675983429
step: 470, loss: 0.01242485549300909
step: 480, loss: 0.008889163844287395
step: 490, loss: 0.0007406350923702121
step: 500, loss: 0.01812712289392948
step: 510, loss: 0.00975298136472702
step: 520, loss: 0.02450649067759514
step: 530, loss: 0.16757068037986755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9435520881138136, f1=0.9381584974805315,
best_f1=0.9381584974805315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035750395618379116
step: 10, loss: 0.006587430834770203
step: 20, loss: 0.004776561167091131
step: 30, loss: 0.11656589806079865
step: 40, loss: 0.0006336601218208671
step: 50, loss: 0.08282753080129623
step: 60, loss: 0.030801953747868538
step: 70, loss: 0.0205239225178957
step: 80, loss: 0.004209682811051607
step: 90, loss: 0.013603412546217442
step: 100, loss: 0.04547962546348572
step: 110, loss: 0.004081699065864086
step: 120, loss: 0.018397115170955658
step: 130, loss: 0.015462375245988369
step: 140, loss: 0.12633658945560455
step: 150, loss: 0.0330280177295208
step: 160, loss: 0.14740432798862457
step: 170, loss: 0.01873987354338169
step: 180, loss: 0.006296120118349791
step: 190, loss: 0.013902561739087105
step: 200, loss: 0.004534165374934673
step: 210, loss: 0.0010517020709812641
step: 220, loss: 0.055157266557216644
step: 230, loss: 0.0016158267389982939
step: 240, loss: 0.0006463438039645553

```

```

step: 250, loss: 0.06653736531734467
step: 260, loss: 0.001284392667002976
step: 270, loss: 0.00030500133289024234
step: 280, loss: 0.005616534035652876
step: 290, loss: 0.11389440298080444
step: 300, loss: 0.2420751303434372
step: 310, loss: 0.022187653928995132
step: 320, loss: 0.03524736687541008
step: 330, loss: 0.0017561016138643026
step: 340, loss: 0.014573248103260994
step: 350, loss: 0.0025310323107987642
step: 360, loss: 0.0007457950268872082
step: 370, loss: 0.0029679681174457073
step: 380, loss: 0.0006992437411099672
step: 390, loss: 0.01851268671452999
step: 400, loss: 0.007972542196512222
step: 410, loss: 0.12417398393154144
step: 420, loss: 0.12061924487352371
step: 430, loss: 0.010016914457082748
step: 440, loss: 0.004499513655900955
step: 450, loss: 0.007525395601987839
step: 460, loss: 0.012149220332503319
step: 470, loss: 0.07965866476297379
step: 480, loss: 0.039920832961797714
step: 490, loss: 0.011118725873529911
step: 500, loss: 0.04110163077712059
step: 510, loss: 0.010228041559457779
step: 520, loss: 0.14448437094688416
step: 530, loss: 0.007748476229608059
/usr/local/lib/python3.7/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=0.941398865784499, f1=0.9304099142040038,
best_f1=0.9381584974805315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019375726580619812
step: 10, loss: 0.0021322753746062517
step: 20, loss: 0.0007734172395430505
step: 30, loss: 0.0009156115702353418
step: 40, loss: 0.0007636764785274863
step: 50, loss: 0.003734809346497059
step: 60, loss: 0.07622221857309341
step: 70, loss: 0.0006368575268425047
step: 80, loss: 0.0024883225560188293
step: 90, loss: 0.009627173654735088
step: 100, loss: 0.04460592567920685

```

```
step: 110, loss: 0.11761917918920517
step: 120, loss: 0.023484988138079643
step: 130, loss: 0.0037981036584824324
step: 140, loss: 0.000604087021201849
step: 150, loss: 0.00044359613093547523
step: 160, loss: 0.04602143540978432
step: 170, loss: 0.0005777888582088053
step: 180, loss: 0.0004532583407126367
step: 190, loss: 0.019437450915575027
step: 200, loss: 0.0015899352729320526
step: 210, loss: 0.04129621013998985
step: 220, loss: 0.0023209687788039446
step: 230, loss: 0.09218860417604446
step: 240, loss: 0.014175716787576675
step: 250, loss: 0.025168580934405327
step: 260, loss: 0.003491007024422288
step: 270, loss: 0.0004085190303158015
step: 280, loss: 0.0011993465013802052
step: 290, loss: 0.005569757893681526
step: 300, loss: 0.0001382537593599409
step: 310, loss: 0.03796852380037308
step: 320, loss: 0.0005536065436899662
step: 330, loss: 0.004331656731665134
step: 340, loss: 0.00020185572793707252
step: 350, loss: 0.11252565681934357
step: 360, loss: 0.09037507325410843
step: 370, loss: 0.003978431690484285
step: 380, loss: 0.0005435055936686695
step: 390, loss: 0.0013351705856621265
step: 400, loss: 0.02009001187980175
step: 410, loss: 0.00022262113634496927
step: 420, loss: 0.03019781969487667
step: 430, loss: 0.06437674164772034
step: 440, loss: 0.005674468353390694
step: 450, loss: 0.09944982081651688
step: 460, loss: 0.01653546839952469
step: 470, loss: 0.01130759622901678
step: 480, loss: 0.0067077321000397205
step: 490, loss: 0.007827279157936573
step: 500, loss: 0.017328625544905663
step: 510, loss: 0.08063396066427231
step: 520, loss: 0.0034638920333236456
step: 530, loss: 0.0008203488541767001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9469732519943688, f1=0.9349056603773584,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022951823193579912
step: 10, loss: 0.003283457364886999
step: 20, loss: 0.0011722323251888156
step: 30, loss: 0.006394416093826294
step: 40, loss: 0.002151346765458584
step: 50, loss: 0.0015474653337150812
step: 60, loss: 0.0009715870837680995
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.00162588304374367
step: 80, loss: 0.0014920373214408755
step: 90, loss: 0.00023176574904937297
step: 100, loss: 0.020166179165244102
step: 110, loss: 0.0004002487112302333
step: 120, loss: 0.001066242577508092
step: 130, loss: 0.0011929728789255023
step: 140, loss: 0.0009754560305736959
step: 150, loss: 0.003947054035961628
step: 160, loss: 0.010311502031981945
step: 170, loss: 0.0011318738106638193
step: 180, loss: 0.11403719335794449
step: 190, loss: 0.010104517452418804
step: 200, loss: 0.0008178330026566982
step: 210, loss: 0.0015616320306435227
step: 220, loss: 0.003531190101057291
step: 230, loss: 0.014229471795260906
step: 240, loss: 0.002072344534099102
step: 250, loss: 0.0325096994638443
step: 260, loss: 0.005386862438172102
step: 270, loss: 0.014735192060470581
step: 280, loss: 0.006639399100095034
step: 290, loss: 0.0020855432376265526
step: 300, loss: 0.0012343103298917413
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 310, loss: 0.0005506825400516391
step: 320, loss: 0.003023984143510461
step: 330, loss: 0.0002938846591860056
step: 340, loss: 0.0001984932750929147
step: 350, loss: 0.0009669850114732981
step: 360, loss: 0.04469488933682442
step: 370, loss: 0.0016711204079911113
step: 380, loss: 0.005711082369089127

```

```

step: 390, loss: 0.00157967081759125
step: 400, loss: 0.05181380361318588
step: 410, loss: 0.00036338617792353034
step: 420, loss: 0.02304176054894924
step: 430, loss: 0.0003309622989036143
step: 440, loss: 0.0005740330670960248
step: 450, loss: 0.0014076497172936797
step: 460, loss: 0.0014329825062304735
step: 470, loss: 0.15202245116233826
step: 480, loss: 0.00953642651438713
step: 490, loss: 0.001877151313237846
step: 500, loss: 0.0005698169698007405
step: 510, loss: 0.0012009850470349193
step: 520, loss: 0.0005035706562921405
step: 530, loss: 0.0007949876016937196
/usr/local/lib/python3.7/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=0.9446768944676894, f1=0.9405756731662025,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008074013749137521
step: 10, loss: 0.0005828975699841976
step: 20, loss: 0.00018563377670943737
step: 30, loss: 0.0009526030626147985
step: 40, loss: 0.00013125990517437458
step: 50, loss: 0.0008538402616977692
step: 60, loss: 0.0006501636817120016
step: 70, loss: 0.0007013324648141861
step: 80, loss: 0.00039537681732326746
step: 90, loss: 0.0003385785676073283
step: 100, loss: 0.0023214297834783792
step: 110, loss: 0.0005204894696362317
step: 120, loss: 0.0005571102956309915
step: 130, loss: 0.1406281739473343
step: 140, loss: 0.0007438297616317868
step: 150, loss: 0.013069374486804008
step: 160, loss: 0.0012518413132056594
step: 170, loss: 0.028075525537133217
step: 180, loss: 0.0004150441091042012
step: 190, loss: 0.005078452173620462
step: 200, loss: 0.0035423170775175095
step: 210, loss: 0.07422547787427902
step: 220, loss: 0.005156909115612507
step: 230, loss: 0.023287395015358925
step: 240, loss: 0.005965664051473141

```

```

step: 250, loss: 0.00012537476141005754
step: 260, loss: 9.926060738507658e-05
step: 270, loss: 0.0009271868038922548
step: 280, loss: 9.258488717023283e-05
step: 290, loss: 0.0016562787350267172
step: 300, loss: 0.00014187530905473977
step: 310, loss: 0.00017020954692270607
step: 320, loss: 0.00013449802645482123
step: 330, loss: 0.002313310047611594
step: 340, loss: 0.008186622522771358
step: 350, loss: 0.000520023109856993
step: 360, loss: 0.0019489596597850323
step: 370, loss: 0.05286982282996178
step: 380, loss: 5.2861574658891186e-05
step: 390, loss: 0.0011234107660129666
step: 400, loss: 0.005294386297464371
step: 410, loss: 0.011118893511593342
step: 420, loss: 0.0016798132564872503
step: 430, loss: 0.0007561521488241851
step: 440, loss: 0.0004789613012690097
step: 450, loss: 0.01221480593085289
step: 460, loss: 0.0009430744685232639
step: 470, loss: 0.13660956919193268
step: 480, loss: 0.004962082952260971
step: 490, loss: 0.000967266212683171
step: 500, loss: 0.003758104983717203
step: 510, loss: 0.01618349179625511
step: 520, loss: 0.008011155761778355
step: 530, loss: 0.0003693084290716797
/usr/local/lib/python3.7/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=0.9442622950819672, f1=0.9445738239403819,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.85759844398126e-05
step: 10, loss: 0.0017742188647389412
step: 20, loss: 0.0026518143713474274
step: 30, loss: 0.01029957365244627
step: 40, loss: 0.0003252739261370152
step: 50, loss: 0.0002865823917090893
step: 60, loss: 0.0005979099660180509
step: 70, loss: 0.0012978892773389816
step: 80, loss: 0.000354442629031837
step: 90, loss: 0.14537163078784943
step: 100, loss: 0.00022988115961197764

```

```
step: 110, loss: 0.005357678048312664
step: 120, loss: 0.014576774090528488
step: 130, loss: 0.00040068122325465083
step: 140, loss: 0.001147893606685102
step: 150, loss: 0.0011489827884361148
step: 160, loss: 0.0013351809466257691
step: 170, loss: 0.0010012483689934015
step: 180, loss: 0.0002079713303828612
step: 190, loss: 0.00010603787086438388
step: 200, loss: 0.00034579477505758405
step: 210, loss: 5.369084465201013e-05
step: 220, loss: 0.00014390995784197003
step: 230, loss: 0.0001530032022856176
step: 240, loss: 9.685365512268618e-05
step: 250, loss: 0.01039244793355465
step: 260, loss: 7.40473042242229e-05
step: 270, loss: 0.0004792460531461984
step: 280, loss: 0.006805530283600092
step: 290, loss: 0.00011896157957380638
step: 300, loss: 0.04910878464579582
step: 310, loss: 0.018170472234487534
step: 320, loss: 8.267030352726579e-05
step: 330, loss: 0.0008856563363224268
step: 340, loss: 0.0067295110784471035
step: 350, loss: 0.045233793556690216
step: 360, loss: 0.000887158268596977
step: 370, loss: 0.0031365747563540936
step: 380, loss: 0.002576689701527357
step: 390, loss: 0.0006798096583224833
step: 400, loss: 0.0014202450402081013
step: 410, loss: 0.00016313919331878424
step: 420, loss: 6.118620513007045e-05
step: 430, loss: 0.0003235993208363652
step: 440, loss: 0.0011444415431469679
step: 450, loss: 0.00024913219385780394
step: 460, loss: 0.003137675579637289
step: 470, loss: 0.00011266575165791437
step: 480, loss: 0.00020785219385288656
step: 490, loss: 0.0010413851123303175
step: 500, loss: 0.026849741116166115
step: 510, loss: 0.00840597040951252
step: 520, loss: 0.005092864856123924
step: 530, loss: 0.01166278775781393
/usr/local/lib/python3.7/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=0.9389454209065681, f1=0.939435968562182,
best_f1=0.9349056603773584
```

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

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

step: 0, loss: 0.0010258551919832826
step: 10, loss: 0.0012292272876948118
step: 20, loss: 0.0007010842673480511
step: 30, loss: 0.0009555937140248716
step: 40, loss: 8.535337110515684e-05
step: 50, loss: 0.00010555362678132951
step: 60, loss: 0.00037568568950518966
step: 70, loss: 0.0009707208373583853
step: 80, loss: 0.0003329823666717857
step: 90, loss: 0.006014854181557894
step: 100, loss: 0.00043833747622556984
step: 110, loss: 0.0527804009616375
step: 120, loss: 0.00020758538448717445
step: 130, loss: 5.2476600103545934e-05
step: 140, loss: 0.07068726420402527
step: 150, loss: 0.002090168884024024
step: 160, loss: 0.07281368225812912
step: 170, loss: 0.00015968368097674102
step: 180, loss: 0.007735051680356264
step: 190, loss: 0.0007553872419521213
step: 200, loss: 0.006690369453281164
step: 210, loss: 0.009760584682226181
step: 220, loss: 0.010207279585301876
step: 230, loss: 0.00036249010008759797
step: 240, loss: 0.00016826295177452266
step: 250, loss: 0.0002444725832901895
step: 260, loss: 0.0025634546764194965
step: 270, loss: 0.0009833131916821003
step: 280, loss: 0.0014133756048977375
step: 290, loss: 0.0006870808429084718
step: 300, loss: 0.00044472457375377417
step: 310, loss: 0.009219101630151272
step: 320, loss: 0.0016470798291265965
step: 330, loss: 0.00016573733591940254
step: 340, loss: 0.0001464730012230575
step: 350, loss: 0.00033389442251063883
step: 360, loss: 0.0015510984230786562
step: 370, loss: 0.0001263450540136546
step: 380, loss: 0.00324221677146852
step: 390, loss: 0.00012143250205554068
step: 400, loss: 0.0015958846779540181
step: 410, loss: 0.0023149331100285053
step: 420, loss: 0.0005416039493866265
step: 430, loss: 0.0004830931138712913
step: 440, loss: 0.0005259810714051127


```

step: 450, loss: 0.07222497463226318
step: 460, loss: 0.00029563758289441466
step: 470, loss: 0.0019327644258737564
step: 480, loss: 0.0006021482404321432
step: 490, loss: 0.0016799404984340072
step: 500, loss: 0.014876622706651688
step: 510, loss: 0.0017238362925127149
step: 520, loss: 0.00040239898953586817
step: 530, loss: 0.0013213350903242826
/usr/local/lib/python3.7/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=0.9425925925925925, f1=0.9426456984273821,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017745829245541245
step: 10, loss: 0.0005144384922459722
step: 20, loss: 0.00010711142385844141
step: 30, loss: 0.00038703440804965794
step: 40, loss: 0.00015435584646183997
step: 50, loss: 9.560512262396514e-05
step: 60, loss: 0.0007810565293766558
step: 70, loss: 0.00015109068772289902
step: 80, loss: 0.0001247784384759143
step: 90, loss: 0.00020709734235424548
step: 100, loss: 0.00031960205524228513
step: 110, loss: 7.28080267435871e-05
step: 120, loss: 0.00014255999121814966
step: 130, loss: 9.557758312439546e-05
step: 140, loss: 9.894457616610453e-05
step: 150, loss: 7.730665674898773e-05
step: 160, loss: 0.0003765807778108865
step: 170, loss: 0.023263122886419296
step: 180, loss: 0.0004930510185658932
step: 190, loss: 0.019089212641119957
step: 200, loss: 0.0002086093445541337
step: 210, loss: 0.00019389994849916548
step: 220, loss: 0.005722166039049625
step: 230, loss: 7.203301356639713e-05
step: 240, loss: 0.0014558180700987577
step: 250, loss: 0.000480032671475783
step: 260, loss: 0.0019442617194727063
step: 270, loss: 0.00044384150533005595
step: 280, loss: 0.0007351734093390405
step: 290, loss: 0.0002643921470735222
step: 300, loss: 9.365848382003605e-05

```

```

step: 310, loss: 0.0009988014353439212
step: 320, loss: 0.0015217813197523355
step: 330, loss: 8.700884791323915e-05
step: 340, loss: 0.11533299833536148
step: 350, loss: 7.009836554061621e-05
step: 360, loss: 0.00029289518715813756
step: 370, loss: 0.0021927254274487495
step: 380, loss: 0.0001712294324534014
step: 390, loss: 0.0010164751438423991
step: 400, loss: 0.026158366352319717
step: 410, loss: 0.00029305496718734503
step: 420, loss: 3.638491762103513e-05
step: 430, loss: 0.00018485479813534766
step: 440, loss: 0.00015224474191199988
step: 450, loss: 0.0002918141835834831
step: 460, loss: 0.000223134396946989
step: 470, loss: 0.0010830911342054605
step: 480, loss: 4.960785008734092e-05
step: 490, loss: 6.059497536625713e-05
step: 500, loss: 0.00013653522182721645
step: 510, loss: 0.0004225037700962275
step: 520, loss: 0.013249440118670464
step: 530, loss: 0.04097043350338936
/usr/local/lib/python3.7/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=0.9359513791491351, f1=0.9351376574895007,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023170428175944835
step: 10, loss: 0.0002287225506734103
step: 20, loss: 0.00029901365633122623
step: 30, loss: 0.00014075216313358396
step: 40, loss: 0.00011913319758605212
step: 50, loss: 0.00018640021153260022
step: 60, loss: 0.00012201490608276799
step: 70, loss: 0.0012128455564379692
step: 80, loss: 9.161361958831549e-05
step: 90, loss: 0.0005598866264335811
step: 100, loss: 0.09359641373157501
step: 110, loss: 0.00010552961612120271
step: 120, loss: 0.00039391263271681964
step: 130, loss: 0.0007750345976091921
step: 140, loss: 9.525638597551733e-05
step: 150, loss: 5.960324051557109e-05
step: 160, loss: 0.003696249797940254

```

```

step: 170, loss: 0.03262777999043465
step: 180, loss: 3.554469731170684e-05
step: 190, loss: 0.00013748272613156587
step: 200, loss: 0.0014440905069932342
step: 210, loss: 0.0007932828739285469
step: 220, loss: 0.0011151123326271772
step: 230, loss: 5.1240596803836524e-05
step: 240, loss: 0.0008175773546099663
step: 250, loss: 0.00023721106117591262
step: 260, loss: 0.00032817345345392823
step: 270, loss: 0.0020299917086958885
step: 280, loss: 0.003653622465208173
step: 290, loss: 0.014420661143958569
step: 300, loss: 0.0025479665491729975
step: 310, loss: 0.00037151548895053566
step: 320, loss: 0.0002443858829792589
step: 330, loss: 0.0011398589704185724
step: 340, loss: 0.0009373481734655797
step: 350, loss: 0.0003276806091889739
step: 360, loss: 4.0921851905295625e-05
step: 370, loss: 4.077791891177185e-05
step: 380, loss: 0.0006939903250895441
step: 390, loss: 6.899147410877049e-05
step: 400, loss: 2.7859785404871218e-05
step: 410, loss: 4.596262442646548e-05
step: 420, loss: 3.656030457932502e-05
step: 430, loss: 0.00036197772715240717
step: 440, loss: 0.0016558360075578094
step: 450, loss: 0.0012253409950062633
step: 460, loss: 0.00017914059571921825
step: 470, loss: 0.0030587457586079836
step: 480, loss: 0.00011690473183989525
step: 490, loss: 3.699785884236917e-05
step: 500, loss: 3.8981615944067016e-05
step: 510, loss: 0.002445126883685589
step: 520, loss: 0.00025866381474770606
step: 530, loss: 0.00012269987200852484
/usr/local/lib/python3.7/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=0.9410664172123478, f1=0.941340782122905,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7487631086842157e-05
step: 10, loss: 0.0001685241877567023
step: 20, loss: 8.402499952353537e-05

```

step: 30, loss: 2.396382660663221e-05
step: 40, loss: 7.019157783361152e-05
step: 50, loss: 0.001656301086768508
step: 60, loss: 2.499538823030889e-05
step: 70, loss: 2.5662509870016947e-05
step: 80, loss: 5.417612919700332e-05
step: 90, loss: 2.4466899049002677e-05
step: 100, loss: 2.4954710170277394e-05
step: 110, loss: 1.6145011613843963e-05
step: 120, loss: 2.115905226673931e-05
step: 130, loss: 1.71320243680384e-05
step: 140, loss: 0.00043174318852834404
step: 150, loss: 0.0011430337326601148
step: 160, loss: 5.490351395565085e-05
step: 170, loss: 0.00013152467727195472
step: 180, loss: 4.940558937960304e-05
step: 190, loss: 6.550050602527335e-05
step: 200, loss: 6.628937990171835e-05
step: 210, loss: 0.00016040085756685585
step: 220, loss: 2.208277874160558e-05
step: 230, loss: 4.2810886952793226e-05
step: 240, loss: 0.002094114199280739
step: 250, loss: 0.00060006242711097
step: 260, loss: 0.00034012473770417273
step: 270, loss: 0.00029183467268012464
step: 280, loss: 0.0006328527233563364
step: 290, loss: 1.8178705431637354e-05
step: 300, loss: 2.457487971696537e-05
step: 310, loss: 0.0006606641691178083
step: 320, loss: 0.00034378902637399733
step: 330, loss: 3.103054405073635e-05
step: 340, loss: 0.01913956180214882
step: 350, loss: 1.7470687453169376e-05
step: 360, loss: 0.12663646042346954
step: 370, loss: 0.0017442145617678761
step: 380, loss: 7.687830657232553e-05
step: 390, loss: 1.8931401427835226e-05
step: 400, loss: 9.301312820753083e-05
step: 410, loss: 3.1648822186980397e-05
step: 420, loss: 0.02831992506980896
step: 430, loss: 0.0003348551981616765
step: 440, loss: 2.1933323296252638e-05
step: 450, loss: 2.6105684810318053e-05
step: 460, loss: 0.0020548247266560793
step: 470, loss: 0.00046473636757582426
step: 480, loss: 3.4867040085373446e-05
step: 490, loss: 0.014896083623170853
step: 500, loss: 0.001015359885059297

```

step: 510, loss: 2.4585086066508666e-05
step: 520, loss: 0.0009029143839143217
step: 530, loss: 0.0015540682943537831
/usr/local/lib/python3.7/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=0.9388322520852641, f1=0.9383402874362541,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046047638170421124
step: 10, loss: 0.00019958146731369197
step: 20, loss: 3.184637171216309e-05
step: 30, loss: 2.4179893443943e-05
step: 40, loss: 4.2632833356037736e-05
step: 50, loss: 0.0004529950674623251
step: 60, loss: 0.00011197704588994384
step: 70, loss: 3.25531218550168e-05
step: 80, loss: 0.0001566633873153478
step: 90, loss: 4.65495468233712e-05
step: 100, loss: 3.793099313043058e-05
step: 110, loss: 9.546164801577106e-05
step: 120, loss: 2.0104856957914308e-05
step: 130, loss: 2.056301491393242e-05
step: 140, loss: 0.00012888561468571424
step: 150, loss: 0.0019281684653833508
step: 160, loss: 0.002372296992689371
step: 170, loss: 0.00022291587083600461
step: 180, loss: 3.221390215912834e-05
step: 190, loss: 0.00014620112779084593
step: 200, loss: 2.4303064492414705e-05
step: 210, loss: 0.006289522163569927
step: 220, loss: 1.4874782209517434e-05
step: 230, loss: 3.053574255318381e-05
step: 240, loss: 1.3880106962460559e-05
step: 250, loss: 0.0005396714550442994
step: 260, loss: 0.0030867597088217735
step: 270, loss: 2.4928629500209354e-05
step: 280, loss: 1.84878572326852e-05
step: 290, loss: 2.1541483874898404e-05
step: 300, loss: 2.0522056729532778e-05
step: 310, loss: 7.318217103602365e-05
step: 320, loss: 2.2291362256510183e-05
step: 330, loss: 0.00046677791397087276
step: 340, loss: 0.00029074324993416667
step: 350, loss: 5.6504610256524757e-05
step: 360, loss: 0.0012872450752183795

```

```

step: 370, loss: 0.00015852742944844067
step: 380, loss: 0.0007644193829037249
step: 390, loss: 1.662511203903705e-05
step: 400, loss: 0.00012771040201187134
step: 410, loss: 1.7343592844554223e-05
step: 420, loss: 2.2744769012206234e-05
step: 430, loss: 7.082570664351806e-05
step: 440, loss: 3.438204657868482e-05
step: 450, loss: 1.3787068382953294e-05
step: 460, loss: 0.0023833459708839655
step: 470, loss: 2.1248444681987166e-05
step: 480, loss: 1.3641784789797384e-05
step: 490, loss: 0.0010238102404400706
step: 500, loss: 1.4632652892032638e-05
step: 510, loss: 0.0016121462685987353
step: 520, loss: 9.711715392768383e-06
step: 530, loss: 4.213667125441134e-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=0.9408476944573824, f1=0.942271880819367,
best_f1=0.9349056603773584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.426377250230871e-05
step: 10, loss: 3.257780190324411e-05
step: 20, loss: 0.0002996872935909778
step: 30, loss: 0.00014720289618708193
step: 40, loss: 2.1822059352416545e-05
step: 50, loss: 0.021204140037298203
step: 60, loss: 0.0027819748502224684
step: 70, loss: 6.291174940997735e-05
step: 80, loss: 1.3544922694563866e-05
step: 90, loss: 0.00010353734978707507
step: 100, loss: 1.1864710359077435e-05
step: 110, loss: 0.00011901794641744345
step: 120, loss: 1.5921401427476667e-05
step: 130, loss: 1.2263468306628056e-05
step: 140, loss: 1.175685156340478e-05
step: 150, loss: 9.827270696405321e-05
step: 160, loss: 1.3485348063113634e-05
step: 170, loss: 1.1648777217487805e-05
step: 180, loss: 5.6241886341013014e-05
step: 190, loss: 0.0038726634811609983
step: 200, loss: 0.008528542704880238
step: 210, loss: 1.0814374945766758e-05
step: 220, loss: 0.00029650144279003143

```

```

step: 230, loss: 1.6361193047487177e-05
step: 240, loss: 1.5224919479805976e-05
step: 250, loss: 1.0918667612713762e-05
step: 260, loss: 1.3574705008068122e-05
step: 270, loss: 2.24250561586814e-05
step: 280, loss: 1.2572667401400395e-05
step: 290, loss: 0.0033377446234226227
step: 300, loss: 1.4360687600856181e-05
step: 310, loss: 0.0011036681244149804
step: 320, loss: 2.2339369024848565e-05
step: 330, loss: 2.2894497305969708e-05
step: 340, loss: 3.858069612761028e-05
step: 350, loss: 0.001311088097281754
step: 360, loss: 1.7623558960622177e-05
step: 370, loss: 0.0003321195545140654
step: 380, loss: 1.4904585441399831e-05
step: 390, loss: 0.0012168745743110776
step: 400, loss: 0.00037797720870003104
step: 410, loss: 2.659964593476616e-05
step: 420, loss: 3.258927245042287e-05
step: 430, loss: 1.0259313057758845e-05
step: 440, loss: 1.4855936569801997e-05
step: 450, loss: 0.000741818337701261
step: 460, loss: 2.592638338683173e-05
step: 470, loss: 1.2684441571764182e-05
step: 480, loss: 0.002031102776527405
step: 490, loss: 1.2822252756450325e-05
step: 500, loss: 4.407468077261001e-05
step: 510, loss: 1.3969472092867363e-05
step: 520, loss: 1.5251019249262754e-05
step: 530, loss: 1.2680689906119369e-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=0.9405756731662025, f1=0.9439555349698935,
best_f1=0.9349056603773584
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, 256.52it/s]
load_f1 = 0.9486461251167133
real_f1 = 0.9460853258321612
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 203.83it/s]

```

3 DK

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                                     | 827.9 MB 1.3 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: 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: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from

```



```

spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.6)
Requirement already satisfied: srsly<1.1.0,>=0.1.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: 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: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: 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: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: 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: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from

```

```

requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=a7b87dc8876d20f97e29cbd5e8ca03df1670243ca6825360629adccc09dd2da7
  Stored in directory: /tmp/pip-ephem-wheel-cache-6xpxni51/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

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

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

# Step 3: run the matcher

```

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

Selected optimization level 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: 0.46710237860679626
```

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=0.288659793814433, f1=0.27368421052631575,
best_f1=0.27368421052631575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39355117082595825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3111111111111111, f1=0.29885057471264365,
best_f1=0.29885057471264365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3992392122745514
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4324324324324324, f1=0.4210526315789474,
best_f1=0.4210526315789474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24722087383270264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4444444444444444, 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: 0.2444545328617096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.5806451612903226, f1=0.5454545454545454,
best_f1=0.5454545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1592542976140976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7200000000000001, f1=0.6000000000000001,
best_f1=0.6000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3187611401081085
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8, f1=0.6428571428571429, best_f1=0.6428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3075312077999115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.7647058823529412,
best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09814007580280304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.75, f1=0.5925925925925927, best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09493247419595718
/usr/local/lib/python3.7/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=0.8, f1=0.6666666666666666, best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19270741939544678
/usr/local/lib/python3.7/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=0.782608695652174, f1=0.6666666666666666,
best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15462400019168854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7999999999999999, f1=0.6000000000000001,
best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035257529467344284
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.6250000000000001,
best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009236043319106102
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.6666666666666667,
best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02268107235431671
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.6666666666666667,
best_f1=0.7647058823529412
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, 114105.13it/s]

```

```

/usr/local/lib/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 = 0.5853658536585367
real_f1 = 0.5581395348837208
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:20, 209.45it/s]
```

3.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:
enabled : True


```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                  : True
opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6069757342338562
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44896355271339417
step: 20, loss: 0.6465490460395813
step: 30, loss: 0.36609944701194763
step: 40, loss: 0.28917303681373596
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.41793903708457947
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 60, loss: 0.1842813789844513
step: 70, loss: 0.12349046766757965
step: 80, loss: 0.09101446717977524
step: 90, loss: 0.2036992609500885
step: 100, loss: 0.1439789980649948
step: 110, loss: 0.2261803299188614
step: 120, loss: 0.1494719386100769
step: 130, loss: 0.02304673194885254
step: 140, loss: 0.035320281982421875
step: 150, loss: 0.13758908212184906
step: 160, loss: 0.07173099368810654
step: 170, loss: 0.04363473877310753
step: 180, loss: 0.15688425302505493
step: 190, loss: 0.04312608391046524
step: 200, loss: 0.03449396416544914
step: 210, loss: 0.07657698541879654
step: 220, loss: 0.1950826346874237

```

```

step: 230, loss: 0.04498426243662834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.968609865470852, f1=0.9646522234891676,
best_f1=0.9646522234891676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054566506296396255
step: 10, loss: 0.12184548377990723
step: 20, loss: 0.02063583955168724
step: 30, loss: 0.013993542641401291
step: 40, loss: 0.015330192632973194
step: 50, loss: 0.07108745723962784
step: 60, loss: 0.011629359796643257
step: 70, loss: 0.014662893489003181
step: 80, loss: 0.01163136400282383
step: 90, loss: 0.04106602072715759
step: 100, loss: 0.004693430848419666
step: 110, loss: 0.0235745869576931
step: 120, loss: 0.015720093622803688
step: 130, loss: 0.012154627591371536
step: 140, loss: 0.06252599507570267
step: 150, loss: 0.0170767679810524
step: 160, loss: 0.037795599550008774
step: 170, loss: 0.0007181353284977376
step: 180, loss: 0.0026460732333362103
step: 190, loss: 0.05939202755689621
step: 200, loss: 0.007587258704006672
step: 210, loss: 0.008796336129307747
step: 220, loss: 0.0006749610183760524
step: 230, loss: 0.0014861897798255086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9775280898876404, f1=0.9737142857142858,
best_f1=0.9737142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029883901588618755
step: 10, loss: 0.07695452868938446
step: 20, loss: 0.08569004386663437
step: 30, loss: 0.017239471897482872
step: 40, loss: 0.09127725660800934
step: 50, loss: 0.06733503937721252
step: 60, loss: 0.0015019197016954422
step: 70, loss: 0.016254903748631477
step: 80, loss: 0.0032725443597882986
step: 90, loss: 0.009462250396609306
step: 100, loss: 0.0026933501940220594
step: 110, loss: 0.047638796269893646
step: 120, loss: 0.0006749049061909318
step: 130, loss: 0.0013138583162799478
step: 140, loss: 0.004444560501724482
step: 150, loss: 0.0047431462444365025
step: 160, loss: 0.022207792848348618
step: 170, loss: 0.01873418316245079
step: 180, loss: 0.01315117534250021
step: 190, loss: 0.016859253868460655
step: 200, loss: 0.001988806063309312
step: 210, loss: 0.004010435193777084
step: 220, loss: 0.006846253294497728
step: 230, loss: 0.012202655896544456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9844097995545658, f1=0.978675645342312,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01712006889283657
step: 10, loss: 0.0010274187661707401
step: 20, loss: 0.0007921489304862916
step: 30, loss: 0.002069918904453516
step: 40, loss: 0.46870386600494385
step: 50, loss: 0.009296941570937634
step: 60, loss: 0.025623461231589317
step: 70, loss: 0.01598389632999897
step: 80, loss: 0.025449179112911224
step: 90, loss: 0.004214255604892969
step: 100, loss: 0.001033923588693142

```

```

step: 110, loss: 0.005777695681899786
step: 120, loss: 0.0006397368269972503
step: 130, loss: 0.005884246435016394
step: 140, loss: 0.0019825794734060764
step: 150, loss: 0.04199974611401558
step: 160, loss: 0.0032823423389345407
step: 170, loss: 0.00321782985702157
step: 180, loss: 0.08819226920604706
step: 190, loss: 0.0014936174266040325
step: 200, loss: 0.0027984334155917168
step: 210, loss: 0.007467618677765131
step: 220, loss: 0.002916014287620783
step: 230, loss: 0.03907586261630058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9799107142857142, f1=0.9798206278026906,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00308394618332386
step: 10, loss: 0.001774700591340661
step: 20, loss: 0.001017103553749621
step: 30, loss: 0.0008471110486425459
step: 40, loss: 0.0008489052415825427
step: 50, loss: 0.0005438889493234456
step: 60, loss: 0.0006927265203557909
step: 70, loss: 0.0011692661792039871
step: 80, loss: 0.13761579990386963
step: 90, loss: 0.22527675330638885
step: 100, loss: 0.0009644034435041249
step: 110, loss: 0.004413447342813015
step: 120, loss: 0.007513085380196571
step: 130, loss: 0.007267479319125414
step: 140, loss: 0.0018764529377222061
step: 150, loss: 0.10762759298086166
step: 160, loss: 0.0025838075671344995
step: 170, loss: 0.001063009025529027
step: 180, loss: 0.001750067574903369
step: 190, loss: 0.04857875779271126
step: 200, loss: 0.0033034654334187508
step: 210, loss: 0.06513724476099014
step: 220, loss: 0.0011231754906475544
step: 230, loss: 0.00849243812263012
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865470852017937, f1=0.9853768278965129,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025615558843128383
step: 10, loss: 0.00430038757622242
step: 20, loss: 0.002721488242968917
step: 30, loss: 0.0006786545272916555
step: 40, loss: 0.0002456969232298434
step: 50, loss: 0.0014822101220488548
step: 60, loss: 0.006817748304456472
step: 70, loss: 0.013626853935420513
step: 80, loss: 0.0035689910873770714
step: 90, loss: 0.006965540815144777
step: 100, loss: 0.0005619805888272822
step: 110, loss: 0.037679340690374374
step: 120, loss: 0.00016322570445481688
step: 130, loss: 0.0006350561161525548
step: 140, loss: 0.004795169923454523
step: 150, loss: 0.0005730746197514236
step: 160, loss: 0.0012957863509654999
step: 170, loss: 0.0015585204819217324
step: 180, loss: 0.00019638962112367153
step: 190, loss: 0.00025894236750900745
step: 200, loss: 0.0006600519409403205
step: 210, loss: 0.00043411582009866834
step: 220, loss: 0.001451970310881734
step: 230, loss: 0.00209772028028965
/usr/local/lib/python3.7/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=0.9819413092550789, f1=0.984090909090909,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005349764134734869
step: 10, loss: 0.0011263178894296288
step: 20, loss: 0.01016014814376831
step: 30, loss: 0.00021402144921012223
step: 40, loss: 0.003313616383820772
step: 50, loss: 0.006282205227762461
step: 60, loss: 0.00024858483811840415

```

```

step: 70, loss: 0.00020445192058105022
step: 80, loss: 0.0005511317867785692
step: 90, loss: 0.014661812223494053
step: 100, loss: 0.001062191789969802
step: 110, loss: 0.00024801495601423085
step: 120, loss: 0.00029308474040590227
step: 130, loss: 0.00016464816872030497
step: 140, loss: 0.00015488517237827182
step: 150, loss: 0.023199457675218582
step: 160, loss: 0.0004573585756588727
step: 170, loss: 0.011317840777337551
step: 180, loss: 0.00026077713118866086
step: 190, loss: 0.00015833371435292065
step: 200, loss: 0.000621286453679204
step: 210, loss: 0.001113726175390184
step: 220, loss: 0.0007400026079267263
step: 230, loss: 0.0018029467901214957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9852774631936579, f1=0.9807037457434733,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011013693874701858
step: 10, loss: 0.0015428692568093538
step: 20, loss: 0.002704452723264694
step: 30, loss: 0.0008836332126520574
step: 40, loss: 0.0013076307950541377
step: 50, loss: 0.00039038920658640563
step: 60, loss: 0.0005581357982009649
step: 70, loss: 0.0002135239337803796
step: 80, loss: 0.036692846566438675
step: 90, loss: 0.00021138580632396042
step: 100, loss: 0.009716175496578217
step: 110, loss: 0.0004469880659598857
step: 120, loss: 0.01611369103193283
step: 130, loss: 0.054091647267341614
step: 140, loss: 0.0003792888019233942
step: 150, loss: 0.009423949755728245
step: 160, loss: 0.000253739912295714
step: 170, loss: 0.03248642757534981
step: 180, loss: 0.00021940236911177635
step: 190, loss: 0.0007092916057445109
step: 200, loss: 0.0034893257543444633
step: 210, loss: 0.000846603128593415
step: 220, loss: 0.0005496809026226401

```

```

step: 230, loss: 0.0012876460095867515
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9784824462061155,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014588718477170914
step: 10, loss: 0.00021155273134354502
step: 20, loss: 0.0006810665363445878
step: 30, loss: 0.008791862055659294
step: 40, loss: 0.00031088810646906495
step: 50, loss: 0.006959256716072559
step: 60, loss: 0.0005635521374642849
step: 70, loss: 0.01263174507766962
step: 80, loss: 0.0006828441983088851
step: 90, loss: 0.004955979064106941
step: 100, loss: 9.137304732576013e-05
step: 110, loss: 9.38407247303985e-05
step: 120, loss: 0.0031978494953364134
step: 130, loss: 0.00011779001215472817
step: 140, loss: 8.272142440546304e-05
step: 150, loss: 0.00018187299428973347
step: 160, loss: 0.009706306271255016
step: 170, loss: 0.0001240638375747949
step: 180, loss: 0.00014173969975672662
step: 190, loss: 0.0006402658764272928
step: 200, loss: 0.001473491545766592
step: 210, loss: 0.07048425078392029
step: 220, loss: 0.00040050665847957134
step: 230, loss: 6.77897478453815e-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=0.9853107344632768, f1=0.9875424688561721,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017714104615151882
step: 10, loss: 0.00026266678469255567
step: 20, loss: 0.0001635297085158527
step: 30, loss: 0.0001239756093127653
step: 40, loss: 0.0014794929884374142
step: 50, loss: 9.567129745846614e-05
step: 60, loss: 0.00023859199427533895

```

```

step: 70, loss: 0.003610338317230344
step: 80, loss: 0.00012947212962899357
step: 90, loss: 0.00019177727517671883
step: 100, loss: 0.00010993570322170854
step: 110, loss: 7.731794903520495e-05
step: 120, loss: 7.17123766662553e-05
step: 130, loss: 0.00010865345393540338
step: 140, loss: 9.445680188946426e-05
step: 150, loss: 0.0001279822608921677
step: 160, loss: 4.423304199008271e-05
step: 170, loss: 8.234294364228845e-05
step: 180, loss: 7.3238777986262e-05
step: 190, loss: 0.0001049483398674056
step: 200, loss: 9.806676825974137e-05
step: 210, loss: 0.0002185964403906837
step: 220, loss: 0.00010432876297272742
step: 230, loss: 0.0001795584976207465
/usr/local/lib/python3.7/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=0.9811738648947952, f1=0.9745293466223698,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.7374122813344e-05
step: 10, loss: 0.00025544711388647556
step: 20, loss: 0.0001439011248294264
step: 30, loss: 0.0001362479815725237
step: 40, loss: 4.1805302316788584e-05
step: 50, loss: 6.00123057665769e-05
step: 60, loss: 0.0009476443519815803
step: 70, loss: 0.20458127558231354
step: 80, loss: 0.006706122774630785
step: 90, loss: 0.07962904870510101
step: 100, loss: 0.0004842159105464816
step: 110, loss: 0.0008661039755679667
step: 120, loss: 0.0004273653030395508
step: 130, loss: 0.00022067478857934475
step: 140, loss: 0.0027285623364150524
step: 150, loss: 0.000240927969571203
step: 160, loss: 0.007136246655136347
step: 170, loss: 0.0001216316013596952
step: 180, loss: 0.00011828169954242185
step: 190, loss: 9.327752195531502e-05
step: 200, loss: 0.0003239052603021264
step: 210, loss: 0.00010462534555699676
step: 220, loss: 0.0005258412566035986

```



```

step: 230, loss: 0.00027307591517455876
/usr/local/lib/python3.7/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=0.9821029082774049, f1=0.9842696629213483,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002548538614064455
step: 10, loss: 7.216101221274585e-05
step: 20, loss: 0.00017068938177544624
step: 30, loss: 0.021541422232985497
step: 40, loss: 0.0001331864477833733
step: 50, loss: 8.631109085399657e-05
step: 60, loss: 8.982946746982634e-05
step: 70, loss: 7.155430648708716e-05
step: 80, loss: 3.617351467255503e-05
step: 90, loss: 0.0003806085733231157
step: 100, loss: 6.657972699031234e-05
step: 110, loss: 5.5481035815319046e-05
step: 120, loss: 0.00038080086233094335
step: 130, loss: 6.808021134929731e-05
step: 140, loss: 0.000839291955344379
step: 150, loss: 0.0002161835291190073
step: 160, loss: 0.0001259925775229931
step: 170, loss: 9.316773503087461e-05
step: 180, loss: 0.0023700541350990534
step: 190, loss: 0.00013782923633698374
step: 200, loss: 5.58027095394209e-05
step: 210, loss: 0.0003559413307812065
step: 220, loss: 2.608296381367836e-05
step: 230, loss: 0.00010337978164898232
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9842696629213483,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.9019909056369215e-05
step: 10, loss: 0.00021381241094786674
step: 20, loss: 8.711338887223974e-05
step: 30, loss: 3.9989943616092205e-05
step: 40, loss: 0.0001406988885719329
step: 50, loss: 0.0011502666166052222
step: 60, loss: 5.271855479804799e-05

```

```

step: 70, loss: 5.8430159697309136e-05
step: 80, loss: 3.931501487386413e-05
step: 90, loss: 3.129851756966673e-05
step: 100, loss: 5.2336297812871635e-05
step: 110, loss: 8.291722042486072e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 120, loss: 9.750732715474442e-05
step: 130, loss: 0.0001049009952112101
step: 140, loss: 4.420510231284425e-05
step: 150, loss: 2.3837152184569277e-05
step: 160, loss: 0.0001607920858077705
step: 170, loss: 3.935206405003555e-05
step: 180, loss: 0.0001601000112714246
step: 190, loss: 0.00014056556392461061
step: 200, loss: 2.7125730412080884e-05
step: 210, loss: 0.00018614988948684186
step: 220, loss: 4.249208723194897e-05
step: 230, loss: 3.5589648177847266e-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=0.9821029082774049, f1=0.9854423292273236,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001960460562258959
step: 10, loss: 3.68886576325167e-05
step: 20, loss: 4.202230047667399e-05
step: 30, loss: 5.778638296760619e-05
step: 40, loss: 3.414810635149479e-05
step: 50, loss: 5.72271783312317e-05
step: 60, loss: 5.163433888810687e-05
step: 70, loss: 5.0367591029498726e-05
step: 80, loss: 3.6661847843788564e-05
step: 90, loss: 8.02908034529537e-05
step: 100, loss: 9.518228762317449e-05
step: 110, loss: 0.00015928370703477412
step: 120, loss: 2.8229302188265137e-05
step: 130, loss: 0.00016693920770194381
step: 140, loss: 0.00010137091157957911
step: 150, loss: 0.0006859332788735628
step: 160, loss: 0.0009132986888289452
step: 170, loss: 0.0012814084766432643
step: 180, loss: 9.755796054378152e-05
step: 190, loss: 5.531738861463964e-05
step: 200, loss: 0.024394461885094643
step: 210, loss: 0.00037557401810772717

```

```

step: 220, loss: 4.3225823901593685e-05
step: 230, loss: 0.1590440720319748
/usr/local/lib/python3.7/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=0.9832402234636871, f1=0.984304932735426,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011303149949526414
step: 10, loss: 6.0808899434050545e-05
step: 20, loss: 0.0005801796214655042
step: 30, loss: 8.504789730068296e-05
step: 40, loss: 5.273225178825669e-05
step: 50, loss: 5.7482153351884335e-05
step: 60, loss: 0.0005887097213417292
step: 70, loss: 0.0005016725626774132
step: 80, loss: 0.0015421538846567273
step: 90, loss: 7.508999988203868e-05
step: 100, loss: 3.5503901017364115e-05
step: 110, loss: 0.00010062693036161363
step: 120, loss: 0.00037706541479565203
step: 130, loss: 5.5926844652276486e-05
step: 140, loss: 0.00012362602865323424
step: 150, loss: 0.00010274922533426434
step: 160, loss: 0.004542328882962465
step: 170, loss: 3.735243808478117e-05
step: 180, loss: 9.675950423115864e-05
step: 190, loss: 0.0008357894839718938
step: 200, loss: 0.00026949975290335715
step: 210, loss: 0.00014764422667212784
step: 220, loss: 0.00011692104453686625
step: 230, loss: 5.746517126681283e-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=0.9832026875699889, f1=0.984304932735426,
best_f1=0.9853768278965129
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, 212.58it/s]
load_f1 = 0.9887640449438202
real_f1 = 0.9843400447427293
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.29it/s]

```

3.1.3 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--dk product \
--summarize

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

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 603kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 35.8MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 20.7MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 70.2MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

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

```

```

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

```

step: 0, loss: 0.6256421208381653
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4306962788105011
step: 20, loss: 0.27634483575820923
step: 30, loss: 0.3282759189605713
step: 40, loss: 0.3888639807701111
step: 50, loss: 0.6874169707298279
step: 60, loss: 0.34216710925102234
step: 70, loss: 0.4655851423740387
step: 80, loss: 0.4805524945259094
step: 90, loss: 0.4611981511116028
step: 100, loss: 0.34017080068588257
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.3066275417804718
step: 120, loss: 0.30816617608070374
step: 130, loss: 0.20225074887275696
step: 140, loss: 0.21523001790046692
step: 150, loss: 0.16157083213329315
step: 160, loss: 0.21323160827159882
step: 170, loss: 0.2531125843524933
step: 180, loss: 0.1260792762041092
step: 190, loss: 0.17511768639087677
step: 200, loss: 0.05736890062689781
step: 210, loss: 0.038421474397182465
step: 220, loss: 0.08292001485824585
step: 230, loss: 0.37419697642326355
step: 240, loss: 0.0936170443892479
step: 250, loss: 0.09388269484043121
step: 260, loss: 0.3744780719280243
step: 270, loss: 0.19522464275360107
step: 280, loss: 0.10287293791770935
step: 290, loss: 0.09722573310136795
step: 300, loss: 0.2209901362657547
step: 310, loss: 0.18596476316452026
step: 320, loss: 0.15802495181560516
step: 330, loss: 0.12385638803243637
step: 340, loss: 0.5989535450935364
step: 350, loss: 0.18916374444961548
step: 360, loss: 0.08770092576742172
step: 370, loss: 0.09506851434707642
step: 380, loss: 0.18394623696804047
step: 390, loss: 0.07205408066511154
step: 400, loss: 0.045692261308431625
step: 410, loss: 0.34358304738998413
step: 420, loss: 0.08496157079935074
step: 430, loss: 0.03635067120194435
step: 440, loss: 0.020094869658350945

```

step: 450, loss: 0.12884582579135895
step: 460, loss: 0.04594554752111435
step: 470, loss: 0.07556210458278656
step: 480, loss: 0.23511847853660583
step: 490, loss: 0.23579354584217072
step: 500, loss: 0.06105966120958328
step: 510, loss: 0.021300414577126503
step: 520, loss: 0.110288105905056
step: 530, loss: 0.18070611357688904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.907871198568873, f1=0.910467706013363,
best_f1=0.910467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06252915412187576
step: 10, loss: 0.06266102939844131
step: 20, loss: 0.15214768052101135
step: 30, loss: 0.1408933848142624
step: 40, loss: 0.11617497354745865
step: 50, loss: 0.11799664795398712
step: 60, loss: 0.034766245633363724
step: 70, loss: 0.06883587688207626
step: 80, loss: 0.07704553753137589
step: 90, loss: 0.16953448951244354
step: 100, loss: 0.18920938670635223
step: 110, loss: 0.04222356155514717
step: 120, loss: 0.24115724861621857
step: 130, loss: 0.031515926122665405
step: 140, loss: 0.05684138461947441
step: 150, loss: 0.012511404231190681
step: 160, loss: 0.0797814354300499
step: 170, loss: 0.12707051634788513
step: 180, loss: 0.15726758539676666
step: 190, loss: 0.013579118065536022
step: 200, loss: 0.23733894526958466
step: 210, loss: 0.08666205406188965
step: 220, loss: 0.005727843381464481
step: 230, loss: 0.028057435527443886
step: 240, loss: 0.05284442380070686
step: 250, loss: 0.03441135585308075
step: 260, loss: 0.19092655181884766

```

```

step: 270, loss: 0.029359672218561172
step: 280, loss: 0.05853551626205444
step: 290, loss: 0.07968654483556747
step: 300, loss: 0.06570262461900711
step: 310, loss: 0.03435707092285156
step: 320, loss: 0.11584748327732086
step: 330, loss: 0.046865370124578476
step: 340, loss: 0.11704444885253906
step: 350, loss: 0.002206384437158704
step: 360, loss: 0.15183177590370178
step: 370, loss: 0.09180520474910736
step: 380, loss: 0.12527213990688324
step: 390, loss: 0.019462864845991135
step: 400, loss: 0.07474105805158615
step: 410, loss: 0.022780891507864
step: 420, loss: 0.022480914369225502
step: 430, loss: 0.46604955196380615
step: 440, loss: 0.1096581444144249
step: 450, loss: 0.0398244634270668
step: 460, loss: 0.029887979850172997
step: 470, loss: 0.1330110877752304
step: 480, loss: 0.032095640897750854
step: 490, loss: 0.08341343700885773
step: 500, loss: 0.047138191759586334
step: 510, loss: 0.036874208599328995
step: 520, loss: 0.28566303849220276
step: 530, loss: 0.06424347311258316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9314942528735632, f1=0.9291628334866605,
best_f1=0.9291628334866605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11237716674804688
step: 10, loss: 0.0592915304005146
step: 20, loss: 0.011807875707745552
step: 30, loss: 0.0651562288403511
step: 40, loss: 0.08225987106561661
step: 50, loss: 0.007285465486347675
step: 60, loss: 0.01581951044499874
step: 70, loss: 0.0036081906873732805
step: 80, loss: 0.006201260723173618

```



```
step: 90, loss: 0.010296694934368134
step: 100, loss: 0.05031628534197807
step: 110, loss: 0.07835887372493744
step: 120, loss: 0.266903817653656
step: 130, loss: 0.0678437203168869
step: 140, loss: 0.007014886476099491
step: 150, loss: 0.046488963067531586
step: 160, loss: 0.01991526037454605
step: 170, loss: 0.0149371437728405
step: 180, loss: 0.017243660986423492
step: 190, loss: 0.021473931148648262
step: 200, loss: 0.042671091854572296
step: 210, loss: 0.18874607980251312
step: 220, loss: 0.1088976114988327
step: 230, loss: 0.02924065850675106
step: 240, loss: 0.06105499714612961
step: 250, loss: 0.1918778121471405
step: 260, loss: 0.040572427213191986
step: 270, loss: 0.03199704363942146
step: 280, loss: 0.062261421233415604
step: 290, loss: 0.03659750148653984
step: 300, loss: 0.16625113785266876
step: 310, loss: 0.03916343301534653
step: 320, loss: 0.006936418823897839
step: 330, loss: 0.1506386697292328
step: 340, loss: 0.02116597257554531
step: 350, loss: 0.04771615192294121
step: 360, loss: 0.023447472602128983
step: 370, loss: 0.065815769135952
step: 380, loss: 0.029948731884360313
step: 390, loss: 0.00919839646667242
step: 400, loss: 0.22018635272979736
step: 410, loss: 0.006260955240577459
step: 420, loss: 0.025757351890206337
step: 430, loss: 0.06732749938964844
step: 440, loss: 0.29356899857521057
step: 450, loss: 0.05916789174079895
step: 460, loss: 0.15273182094097137
step: 470, loss: 0.02973727136850357
step: 480, loss: 0.07632181793451309
step: 490, loss: 0.009464564733207226
step: 500, loss: 0.03557922691106796
step: 510, loss: 0.007843374274671078
step: 520, loss: 0.01257061306387186
step: 530, loss: 0.025665396824479103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9382257315373896, f1=0.9314685314685315,
best_f1=0.9314685314685315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029186246916651726
step: 10, loss: 0.00346185639500618
step: 20, loss: 0.11196126788854599
step: 30, loss: 0.022487295791506767
step: 40, loss: 0.14035820960998535
step: 50, loss: 0.09968649595975876
step: 60, loss: 0.02701244130730629
step: 70, loss: 0.16243121027946472
step: 80, loss: 0.07112375646829605
step: 90, loss: 0.016521258279681206
step: 100, loss: 0.007250742055475712
step: 110, loss: 0.08495263755321503
step: 120, loss: 0.002134740585461259
step: 130, loss: 0.11727315187454224
step: 140, loss: 0.06795845925807953
step: 150, loss: 0.07753583043813705
step: 160, loss: 0.0013037034077569842
step: 170, loss: 0.0857945904135704
step: 180, loss: 0.03566538915038109
step: 190, loss: 0.0699065774679184
step: 200, loss: 0.017551125958561897
step: 210, loss: 0.0028303819708526134
step: 220, loss: 0.01481418963521719
step: 230, loss: 0.00509444298222661
step: 240, loss: 0.012147539295256138
step: 250, loss: 0.08760028332471848
step: 260, loss: 0.00284752668812871
step: 270, loss: 0.09628960490226746
step: 280, loss: 0.01397050078958273
step: 290, loss: 0.02445645071566105
step: 300, loss: 0.0012300560483708978
step: 310, loss: 0.01505535189062357
step: 320, loss: 0.19751417636871338
step: 330, loss: 0.017691290006041527
step: 340, loss: 0.006227446254342794
step: 350, loss: 0.11135277897119522
step: 360, loss: 0.01681319624185562
step: 370, loss: 0.004663281142711639
step: 380, loss: 0.007518659345805645

```

```

step: 390, loss: 0.009901260025799274
step: 400, loss: 0.016122929751873016
step: 410, loss: 0.0035974609199911356
step: 420, loss: 0.008956526406109333
step: 430, loss: 0.006610003765672445
step: 440, loss: 0.006611410528421402
step: 450, loss: 0.05965002626180649
step: 460, loss: 0.07577376812696457
step: 470, loss: 0.017487160861492157
step: 480, loss: 0.005667979829013348
step: 490, loss: 0.0014584045857191086
step: 500, loss: 0.06792086362838745
step: 510, loss: 0.12595976889133453
step: 520, loss: 0.0190895926207304
step: 530, loss: 0.09891132265329361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9409005628517824, f1=0.9359513791491351,
best_f1=0.9359513791491351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010693660005927086
step: 10, loss: 0.006987835746258497
step: 20, loss: 0.0016701689455658197
step: 30, loss: 0.0034406171180307865
step: 40, loss: 0.008726530708372593
step: 50, loss: 0.025824574753642082
step: 60, loss: 0.05041117966175079
step: 70, loss: 0.001528009888716042
step: 80, loss: 0.00032102022669278085
step: 90, loss: 0.032122209668159485
step: 100, loss: 0.041155483573675156
step: 110, loss: 0.018037550151348114
step: 120, loss: 0.2142922580242157
step: 130, loss: 0.005274288821965456
step: 140, loss: 0.07003708183765411
step: 150, loss: 0.0668535828590393
step: 160, loss: 0.02886250987648964
step: 170, loss: 0.04249921813607216
step: 180, loss: 0.023612726479768753
step: 190, loss: 0.0011116339592263103
step: 200, loss: 0.012964689172804356

```

```

step: 210, loss: 0.1769222915172577
step: 220, loss: 0.002782669849693775
step: 230, loss: 0.029877105727791786
step: 240, loss: 0.07088073343038559
step: 250, loss: 0.14727018773555756
step: 260, loss: 0.0225825272500515
step: 270, loss: 0.00331481103785336
step: 280, loss: 0.003673451952636242
step: 290, loss: 0.009815756231546402
step: 300, loss: 0.016082555055618286
step: 310, loss: 0.09578905254602432
step: 320, loss: 0.03630780801177025
step: 330, loss: 0.009712457656860352
step: 340, loss: 0.008864151313900948
step: 350, loss: 0.011126870289444923
step: 360, loss: 0.0006242113304324448
step: 370, loss: 0.011107354424893856
step: 380, loss: 0.004114476498216391
step: 390, loss: 0.003238597186282277
step: 400, loss: 0.005975755397230387
step: 410, loss: 0.04239710792899132
step: 420, loss: 0.24871660768985748
step: 430, loss: 0.08649026602506638
step: 440, loss: 0.008906621485948563
step: 450, loss: 0.0015585562214255333
step: 460, loss: 0.028474798426032066
step: 470, loss: 0.03153960034251213
step: 480, loss: 0.016142241656780243
step: 490, loss: 0.029961103573441505
step: 500, loss: 0.11758997291326523
step: 510, loss: 0.002432950073853135
step: 520, loss: 0.07656639069318771
step: 530, loss: 0.0029919787775725126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9420884632922938, f1=0.9322344322344323,
best_f1=0.9322344322344323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06324082612991333
step: 10, loss: 0.030907798558473587
step: 20, loss: 0.02057468704879284

```

step: 30, loss: 0.02548452839255333
step: 40, loss: 0.0012855560053139925
step: 50, loss: 0.0004144315607845783
step: 60, loss: 0.0003806032764259726
step: 70, loss: 0.0002944486332125962
step: 80, loss: 0.025947904214262962
step: 90, loss: 0.01307179406285286
step: 100, loss: 0.04130158945918083
step: 110, loss: 0.002825275296345353
step: 120, loss: 0.031897809356451035
step: 130, loss: 0.0008800172945484519
step: 140, loss: 0.0072426386177539825
step: 150, loss: 0.0005201312014833093
step: 160, loss: 0.02390635758638382
step: 170, loss: 0.0022536052856594324
step: 180, loss: 0.0010062806541100144
step: 190, loss: 0.07529893517494202
step: 200, loss: 0.036592867225408554
step: 210, loss: 0.0037506050430238247
step: 220, loss: 0.016534458845853806
step: 230, loss: 0.0029056163039058447
step: 240, loss: 0.03429105132818222
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 250, loss: 0.18448424339294434
step: 260, loss: 0.024272693321108818
step: 270, loss: 0.13807237148284912
step: 280, loss: 0.0022197263315320015
step: 290, loss: 0.0007194514619186521
step: 300, loss: 0.0012803103309124708
step: 310, loss: 0.02960415557026863
step: 320, loss: 0.07442639768123627
step: 330, loss: 0.01067386195063591
step: 340, loss: 0.0009339284733869135
step: 350, loss: 0.009552175179123878
step: 360, loss: 0.0943676307797432
step: 370, loss: 0.023649021983146667
step: 380, loss: 0.010776645503938198
step: 390, loss: 0.0014073573984205723
step: 400, loss: 0.021157385781407356
step: 410, loss: 0.00075221509905532
step: 420, loss: 0.003088114084675908
step: 430, loss: 0.0018724487163126469
step: 440, loss: 0.00032097924849949777
step: 450, loss: 0.2612660229206085
step: 460, loss: 0.0018087814096361399
step: 470, loss: 0.022739227861166
step: 480, loss: 0.020516017451882362
step: 490, loss: 0.008314522914588451

```
step: 500, loss: 0.0009098837035708129
step: 510, loss: 0.25038933753967285
step: 520, loss: 0.004733433481305838
step: 530, loss: 0.04415503144264221
/usr/local/lib/python3.7/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=0.9414498141263941, f1=0.9395477618827872,
best_f1=0.9322344322344323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004252455662935972
step: 10, loss: 0.0012124977074563503
step: 20, loss: 0.0019706792663782835
step: 30, loss: 0.0024895165115594864
step: 40, loss: 0.0022454014979302883
step: 50, loss: 0.0020215061958879232
step: 60, loss: 0.0026479954831302166
step: 70, loss: 0.001982814399525523
step: 80, loss: 0.0004132879839744419
step: 90, loss: 0.00028704924625344574
step: 100, loss: 0.0003595990710891783
step: 110, loss: 0.0013991397572681308
step: 120, loss: 0.004318794701248407
step: 130, loss: 0.00021557306172326207
step: 140, loss: 0.008987178094685078
step: 150, loss: 0.0038537613581866026
step: 160, loss: 0.0010940851643681526
step: 170, loss: 0.006955387070775032
step: 180, loss: 0.02646017074584961
step: 190, loss: 0.023398561403155327
step: 200, loss: 0.006372211035341024
step: 210, loss: 0.06744667142629623
step: 220, loss: 0.00029708456713706255
step: 230, loss: 0.0002176321140723303
step: 240, loss: 0.002134912647306919
step: 250, loss: 0.002234329702332616
step: 260, loss: 0.004121270030736923
step: 270, loss: 0.006061078514903784
step: 280, loss: 0.006077141966670752
step: 290, loss: 0.006197589449584484
step: 300, loss: 0.0010375215206295252
step: 310, loss: 0.00867849588394165
step: 320, loss: 0.09347977489233017
step: 330, loss: 0.0011545252054929733
step: 340, loss: 0.00029812095453962684
step: 350, loss: 0.004091854207217693
```

```

step: 360, loss: 0.015496216714382172
step: 370, loss: 0.0002926118322648108
step: 380, loss: 0.0077986884862184525
step: 390, loss: 0.009622813202440739
step: 400, loss: 0.012148255482316017
step: 410, loss: 0.0027116795536130667
step: 420, loss: 0.023658474907279015
step: 430, loss: 0.0024692958686500788
step: 440, loss: 0.0006650149589404464
step: 450, loss: 0.0017860205844044685
step: 460, loss: 0.0012274696491658688
step: 470, loss: 0.25041502714157104
step: 480, loss: 0.0044174096547067165
step: 490, loss: 0.10392676293849945
step: 500, loss: 0.003229991067200899
step: 510, loss: 0.0004002508067060262
step: 520, loss: 0.00041803266503848135
step: 530, loss: 0.0005794028984382749
/usr/local/lib/python3.7/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=0.9420491423273065, f1=0.9404706968158745,
best_f1=0.9322344322344323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030997180147096515
step: 10, loss: 0.0006438489654101431
step: 20, loss: 0.001402486115694046
step: 30, loss: 0.00043240730883553624
step: 40, loss: 0.0012621416244655848
step: 50, loss: 0.0008309410768561065
step: 60, loss: 0.0009708712459541857
step: 70, loss: 0.0009211588767357171
step: 80, loss: 0.003591514192521572
step: 90, loss: 0.0007588383741676807
step: 100, loss: 0.0012568642850965261
step: 110, loss: 0.0012017900589853525
step: 120, loss: 0.003985157236456871
step: 130, loss: 0.0008308584801852703
step: 140, loss: 0.0003640275099314749
step: 150, loss: 0.00011421473027439788
step: 160, loss: 0.00022753779194317758
step: 170, loss: 0.04274697229266167
step: 180, loss: 0.0007241869461722672
step: 190, loss: 0.0002356578770559281
step: 200, loss: 0.00022039670147933066
step: 210, loss: 0.173629030585289

```

```

step: 220, loss: 0.00497534591704607
step: 230, loss: 0.002048656577244401
step: 240, loss: 0.0006717925425618887
step: 250, loss: 0.00020000668882858008
step: 260, loss: 0.001831586705520749
step: 270, loss: 0.00027277227491140366
step: 280, loss: 9.981312177842483e-05
step: 290, loss: 0.003972725477069616
step: 300, loss: 0.0016079485649242997
step: 310, loss: 0.002328320639207959
step: 320, loss: 0.0008466159924864769
step: 330, loss: 0.0002276055165566504
step: 340, loss: 0.015271580778062344
step: 350, loss: 8.259444439318031e-05
step: 360, loss: 0.0003486999776214361
step: 370, loss: 0.0051098717376589775
step: 380, loss: 0.0007693920633755624
step: 390, loss: 0.0015128761297091842
step: 400, loss: 0.000529135693795979
step: 410, loss: 0.0006480145966634154
step: 420, loss: 0.0006245104013942182
step: 430, loss: 0.006628716830164194
step: 440, loss: 0.0062359715811908245
step: 450, loss: 0.001061991206370294
step: 460, loss: 0.03446057438850403
step: 470, loss: 0.060902856290340424
step: 480, loss: 0.0016137979691848159
step: 490, loss: 0.002299276180565357
step: 500, loss: 0.0005297376192174852
step: 510, loss: 0.0009032812668010592
step: 520, loss: 0.0004958354402333498
step: 530, loss: 0.00026509445160627365
/usr/local/lib/python3.7/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=0.9376443418013857, f1=0.9397260273972603,
best_f1=0.9322344322344323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014020642265677452
step: 10, loss: 0.006445765495300293
step: 20, loss: 0.0002998972195200622
step: 30, loss: 0.08514995872974396
step: 40, loss: 0.0004887085524387658
step: 50, loss: 0.0018810027977451682
step: 60, loss: 0.001815838273614645
step: 70, loss: 0.0023448672145605087

```



```
step: 80, loss: 0.11926870048046112
step: 90, loss: 0.003501017577946186
step: 100, loss: 0.029464419931173325
step: 110, loss: 0.01822161115705967
step: 120, loss: 0.0007500117062591016
step: 130, loss: 0.0030768972355872393
step: 140, loss: 0.000280445150565356
step: 150, loss: 0.0029407215770334005
step: 160, loss: 0.06784021854400635
step: 170, loss: 0.003577732015401125
step: 180, loss: 0.006419310346245766
step: 190, loss: 0.00034642068203538656
step: 200, loss: 0.00010263898002449423
step: 210, loss: 9.166354720946401e-05
step: 220, loss: 8.463569247396663e-05
step: 230, loss: 7.742381421849132e-05
step: 240, loss: 0.00020116748055443168
step: 250, loss: 0.005446873139590025
step: 260, loss: 4.669031477533281e-05
step: 270, loss: 0.021471882238984108
step: 280, loss: 0.006701147649437189
step: 290, loss: 0.00193088932428509
step: 300, loss: 0.0002877333608921617
step: 310, loss: 0.00014005396224092692
step: 320, loss: 0.057099781930446625
step: 330, loss: 0.00019939205958507955
step: 340, loss: 0.009179949760437012
step: 350, loss: 0.006551885977387428
step: 360, loss: 9.098400187212974e-05
step: 370, loss: 0.02919859252870083
step: 380, loss: 6.910256342962384e-05
step: 390, loss: 0.00027502261218614876
step: 400, loss: 0.0001003454890451394
step: 410, loss: 0.0017397543415427208
step: 420, loss: 4.7222725697793067e-05
step: 430, loss: 0.00014063685375731438
step: 440, loss: 5.438486550701782e-05
step: 450, loss: 8.730008266866207e-05
step: 460, loss: 5.2619037887779996e-05
step: 470, loss: 4.378874655230902e-05
step: 480, loss: 5.9273657825542614e-05
step: 490, loss: 0.1053772121667862
step: 500, loss: 0.0002067346649710089
step: 510, loss: 0.035311259329319
step: 520, loss: 0.01745995134115219
step: 530, loss: 0.0010406834771856666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9444444444444445, f1=0.9386814200092208,
best_f1=0.9386814200092208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002163020195439458
step: 10, loss: 0.0003744499117601663
step: 20, loss: 0.0007250038324855268
step: 30, loss: 0.0014006636338308454
step: 40, loss: 0.0014911788748577237
step: 50, loss: 0.0005730461562052369
step: 60, loss: 0.018353747203946114
step: 70, loss: 7.405389624182135e-05
step: 80, loss: 6.274252518778667e-05
step: 90, loss: 0.005151588004082441
step: 100, loss: 0.011893248185515404
step: 110, loss: 0.007217321544885635
step: 120, loss: 0.0009596995078027248
step: 130, loss: 0.008806231431663036
step: 140, loss: 0.0014993278309702873
step: 150, loss: 9.567711094859987e-05
step: 160, loss: 0.06006518006324768
step: 170, loss: 0.0002460223622620106
step: 180, loss: 0.0005929350154474378
step: 190, loss: 4.415642615640536e-05
step: 200, loss: 0.0022103306837379932
step: 210, loss: 0.048714954406023026
step: 220, loss: 0.0006403585430234671
step: 230, loss: 0.00042151857633143663
step: 240, loss: 6.540659524034709e-05
step: 250, loss: 0.0016819440061226487
step: 260, loss: 0.0018448979826644063
step: 270, loss: 0.0005839724908582866
step: 280, loss: 0.0025792550295591354
step: 290, loss: 0.0005369577556848526
step: 300, loss: 0.010321966372430325
step: 310, loss: 0.023075900971889496
step: 320, loss: 0.0017856704071164131
step: 330, loss: 0.005075505468994379
step: 340, loss: 0.0009921457385644317
step: 350, loss: 0.042883481830358505
step: 360, loss: 0.0002437916409689933
step: 370, loss: 0.00042597544961608946

```

```

step: 380, loss: 0.0006791057530790567
step: 390, loss: 0.00013489222328644246
step: 400, loss: 0.0007337750284932554
step: 410, loss: 0.024535493925213814
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 420, loss: 0.00013094412861391902
step: 430, loss: 0.003036465961486101
step: 440, loss: 0.00014123560686130077
step: 450, loss: 0.0002149942156393081
step: 460, loss: 6.664084503427148e-05
step: 470, loss: 9.210930875269696e-05
step: 480, loss: 8.150395296979696e-05
step: 490, loss: 0.00013538960774894804
step: 500, loss: 0.00031700488761998713
step: 510, loss: 5.393169340095483e-05
step: 520, loss: 0.0038345158100128174
step: 530, loss: 0.001961065921932459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9456572224802601, f1=0.9422632794457274,
best_f1=0.9422632794457274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020498372032307088
step: 10, loss: 4.4144973799120635e-05
step: 20, loss: 0.00011598120909184217
step: 30, loss: 6.04454216954764e-05
step: 40, loss: 0.0005507160094566643
step: 50, loss: 3.7638623325619847e-05
step: 60, loss: 4.377476579975337e-05
step: 70, loss: 0.00093937071505934
step: 80, loss: 0.010851005092263222
step: 90, loss: 0.00013494386803358793
step: 100, loss: 2.4705524992896244e-05
step: 110, loss: 2.5450624889344908e-05
step: 120, loss: 6.653768650721759e-05
step: 130, loss: 4.5514610974350944e-05
step: 140, loss: 6.461189332185313e-05
step: 150, loss: 0.0002976754039991647
step: 160, loss: 0.0007909185951575637
step: 170, loss: 5.95261954003945e-05
step: 180, loss: 0.00030075464746914804

```

```

step: 190, loss: 0.00128271640278399
step: 200, loss: 0.004911421332508326
step: 210, loss: 0.001401500659994781
step: 220, loss: 0.04619070887565613
step: 230, loss: 2.7588566808844917e-05
step: 240, loss: 0.04710311442613602
step: 250, loss: 0.0017077202210202813
step: 260, loss: 0.00039925798773765564
step: 270, loss: 0.0006521950126625597
step: 280, loss: 0.0004753296379931271
step: 290, loss: 0.0006837928667664528
step: 300, loss: 0.0006023839232511818
step: 310, loss: 0.0036705846432596445
step: 320, loss: 0.00033348111901432276
step: 330, loss: 9.45192514336668e-05
step: 340, loss: 0.0003482752072159201
step: 350, loss: 0.03656958416104317
step: 360, loss: 0.0007461131317541003
step: 370, loss: 0.005492724943906069
step: 380, loss: 0.003124531591311097
step: 390, loss: 0.0002627782814670354
step: 400, loss: 0.008765322156250477
step: 410, loss: 0.00010220703552477062
step: 420, loss: 0.00011137253022752702
step: 430, loss: 0.0023978326935321093
step: 440, loss: 0.000295703619485721
step: 450, loss: 0.00012646075629163533
step: 460, loss: 4.550821540760808e-05
step: 470, loss: 0.0001633259525988251
step: 480, loss: 0.001095529180020094
step: 490, loss: 9.933773253578693e-05
step: 500, loss: 0.0002816544147208333
step: 510, loss: 0.00020757626043632627
step: 520, loss: 0.0007101544179022312
step: 530, loss: 0.011406256817281246
/usr/local/lib/python3.7/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=0.9426573426573427, f1=0.9364858599907279,
best_f1=0.9422632794457274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002567259594798088
step: 10, loss: 0.00046755088260397315
step: 20, loss: 0.004995469003915787
step: 30, loss: 5.7583602028898895e-05
step: 40, loss: 0.0013331101508811116

```

step: 50, loss: 3.34139731421601e-05
step: 60, loss: 4.2067440517712384e-05
step: 70, loss: 0.017028748989105225
step: 80, loss: 0.008068027906119823
step: 90, loss: 0.0002645473286975175
step: 100, loss: 0.15156912803649902
step: 110, loss: 7.260269194375724e-05
step: 120, loss: 0.0006835852982476354
step: 130, loss: 0.00039023952558636665
step: 140, loss: 0.00014861577074043453
step: 150, loss: 0.06428470462560654
step: 160, loss: 0.1841449737548828
step: 170, loss: 7.509938586736098e-05
step: 180, loss: 2.7089337891084142e-05
step: 190, loss: 0.07516659051179886
step: 200, loss: 0.00022203687694855034
step: 210, loss: 0.0019533850718289614
step: 220, loss: 0.003176571801304817
step: 230, loss: 6.218483758857474e-05
step: 240, loss: 0.0031692429911345243
step: 250, loss: 5.01384238305036e-05
step: 260, loss: 0.00044257770059630275
step: 270, loss: 0.003824267303571105
step: 280, loss: 0.0008444084669463336
step: 290, loss: 0.0001860527991084382
step: 300, loss: 0.00043402795563451946
step: 310, loss: 0.004926270339637995
step: 320, loss: 0.0016781764570623636
step: 330, loss: 0.002831026678904891
step: 340, loss: 0.0005226865760050714
step: 350, loss: 0.0007108890567906201
step: 360, loss: 0.00046404398744925857
step: 370, loss: 0.0001335594424745068
step: 380, loss: 0.00011948461178690195
step: 390, loss: 0.007527345325797796
step: 400, loss: 2.1214858861640096e-05
step: 410, loss: 5.372047235141508e-05
step: 420, loss: 0.0013611743925139308
step: 430, loss: 0.00026540979160927236
step: 440, loss: 0.0010452085407450795
step: 450, loss: 0.0002893502824008465
step: 460, loss: 5.540377605939284e-05
step: 470, loss: 0.0025166107807308435
step: 480, loss: 0.00028965939418412745
step: 490, loss: 0.0011857287026941776
step: 500, loss: 1.7784052033675835e-05
step: 510, loss: 6.164271326269954e-05
step: 520, loss: 0.0014193656388670206

```
step: 530, loss: 4.725602047983557e-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=0.9424228466144633, f1=0.9384756657483929,
best_f1=0.9422632794457274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.1326839487301186e-05
step: 10, loss: 1.9992879970232025e-05
step: 20, loss: 3.214762909919955e-05
step: 30, loss: 2.042905725829769e-05
step: 40, loss: 3.467461283435114e-05
step: 50, loss: 0.0007124778348952532
step: 60, loss: 7.545595144620165e-05
step: 70, loss: 0.0018120177555829287
step: 80, loss: 5.072520798421465e-05
step: 90, loss: 0.00012627916294150054
step: 100, loss: 1.8439783161738887e-05
step: 110, loss: 4.852967322221957e-05
step: 120, loss: 1.957956192200072e-05
step: 130, loss: 1.503128805779852e-05
step: 140, loss: 0.0001461694046156481
step: 150, loss: 0.001499262056313455
step: 160, loss: 0.0008358724298886955
step: 170, loss: 4.54615437774919e-05
step: 180, loss: 4.19977332057897e-05
step: 190, loss: 7.522403757320717e-05
step: 200, loss: 2.103985934809316e-05
step: 210, loss: 1.8871864085667767e-05
step: 220, loss: 1.8119459127774462e-05
step: 230, loss: 0.000633353425655514
step: 240, loss: 3.7294277717592195e-05
step: 250, loss: 2.8135951652075164e-05
step: 260, loss: 2.0395371393533424e-05
step: 270, loss: 0.00045152512029744685
step: 280, loss: 0.0002861490356735885
step: 290, loss: 2.548527299950365e-05
step: 300, loss: 2.248523924208712e-05
step: 310, loss: 2.2939342670724727e-05
step: 320, loss: 0.00030239776242524385
step: 330, loss: 9.210612188326195e-05
step: 340, loss: 1.7650114386924542e-05
step: 350, loss: 1.3302789739100263e-05
step: 360, loss: 0.0003433232777751982
step: 370, loss: 0.0001309462677454576
step: 380, loss: 0.00010933924932032824
```

```

step: 390, loss: 1.948989120137412e-05
step: 400, loss: 7.105893746484071e-05
step: 410, loss: 2.3312075427384116e-05
step: 420, loss: 0.00032075404305942357
step: 430, loss: 4.0911756514105946e-05
step: 440, loss: 1.4070194993109908e-05
step: 450, loss: 0.00014848327555228025
step: 460, loss: 5.484242137754336e-05
step: 470, loss: 0.0009499256266281009
step: 480, loss: 1.279987827729201e-05
step: 490, loss: 2.099487028317526e-05
step: 500, loss: 0.0009039246942847967
step: 510, loss: 6.732226756867021e-05
step: 520, loss: 2.0115017832722515e-05
step: 530, loss: 0.07888961583375931
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9471733086190918, f1=0.9443166129774505,
best_f1=0.9443166129774505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.4290475014131516e-05
step: 10, loss: 0.0013111039297655225
step: 20, loss: 1.4137282050796784e-05
step: 30, loss: 6.889519863761961e-05
step: 40, loss: 0.00025462114717811346
step: 50, loss: 0.0010007204255089164
step: 60, loss: 0.00010259906412102282
step: 70, loss: 2.112912989105098e-05
step: 80, loss: 0.003995355684310198
step: 90, loss: 6.772040069336072e-05
step: 100, loss: 1.91773669939721e-05
step: 110, loss: 0.000936525349970907
step: 120, loss: 1.658467408560682e-05
step: 130, loss: 1.592906664882321e-05
step: 140, loss: 0.0009975428692996502
step: 150, loss: 1.4967984498071019e-05
step: 160, loss: 9.181817586068064e-05
step: 170, loss: 0.0109022231772542
step: 180, loss: 2.4048791601671837e-05
step: 190, loss: 5.995376341161318e-05
step: 200, loss: 1.6435735233244486e-05

```

```

step: 210, loss: 0.008874225430190563
step: 220, loss: 2.436192153254524e-05
step: 230, loss: 2.356873665121384e-05
step: 240, loss: 9.477064850216266e-06
step: 250, loss: 4.563877155305818e-05
step: 260, loss: 1.391739169775974e-05
step: 270, loss: 8.236039866460487e-05
step: 280, loss: 7.382596959359944e-05
step: 290, loss: 1.6528143532923423e-05
step: 300, loss: 1.4751924936717842e-05
step: 310, loss: 4.212370913592167e-05
step: 320, loss: 1.714335667202249e-05
step: 330, loss: 0.00022769228962715715
step: 340, loss: 4.7583787818439305e-05
step: 350, loss: 1.3731216313317418e-05
step: 360, loss: 2.8389205908752047e-05
step: 370, loss: 0.00018028159684035927
step: 380, loss: 0.00034308710019104183
step: 390, loss: 1.16897626867285e-05
step: 400, loss: 0.0011224941117689013
step: 410, loss: 1.769430673448369e-05
step: 420, loss: 3.291570828878321e-05
step: 430, loss: 0.00264498102478683
step: 440, loss: 0.0008909771568141878
step: 450, loss: 0.0002071772760245949
step: 460, loss: 0.0002859297674149275
step: 470, loss: 2.9474836992449127e-05
step: 480, loss: 1.3116605259710923e-05
step: 490, loss: 1.5094594346010126e-05
step: 500, loss: 0.0003286896971985698
step: 510, loss: 0.005712212063372135
step: 520, loss: 1.142153087130282e-05
step: 530, loss: 1.3582228348241188e-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=0.9479553903345725, f1=0.9405255878284925,
best_f1=0.9405255878284925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.3351256711757742e-05
step: 10, loss: 0.00022252791677601635
step: 20, loss: 6.838129775132984e-05

```


step: 30, loss: 4.636655285139568e-05
step: 40, loss: 4.447438186616637e-05
step: 50, loss: 0.0024368823505938053
step: 60, loss: 9.307869186159223e-05
step: 70, loss: 3.208907583029941e-05
step: 80, loss: 1.391000478179194e-05
step: 90, loss: 1.9236143998568878e-05
step: 100, loss: 8.899631211534142e-06
step: 110, loss: 0.0003657975175883621
step: 120, loss: 1.217776480189059e-05
step: 130, loss: 0.00085354846669361
step: 140, loss: 9.563037019688636e-05
step: 150, loss: 2.053604657703545e-05
step: 160, loss: 1.0803234545164742e-05
step: 170, loss: 1.2136755685787648e-05
step: 180, loss: 0.0003019215364474803
step: 190, loss: 6.718140502925962e-05
step: 200, loss: 6.484911136794835e-05
step: 210, loss: 1.2740277270495426e-05
step: 220, loss: 1.1484939022921026e-05
step: 230, loss: 0.0001186066510854289
step: 240, loss: 3.313543493277393e-05
step: 250, loss: 1.03487554952153e-05
step: 260, loss: 1.1425334378145635e-05
step: 270, loss: 1.4312345228972845e-05
step: 280, loss: 1.4133439435681794e-05
step: 290, loss: 0.0007366008940152824
step: 300, loss: 2.025675894401502e-05
step: 310, loss: 0.030968671664595604
step: 320, loss: 1.5124431229196489e-05
step: 330, loss: 1.3083076737530064e-05
step: 340, loss: 0.0003649496356956661
step: 350, loss: 0.010177419520914555
step: 360, loss: 0.0003273699840065092
step: 370, loss: 1.7944345017895103e-05
step: 380, loss: 2.318796396139078e-05
step: 390, loss: 1.2933985999552533e-05
step: 400, loss: 0.005561104975640774
step: 410, loss: 3.3114130928879604e-05
step: 420, loss: 6.513233529403806e-05
step: 430, loss: 1.0121455488842912e-05
step: 440, loss: 2.6100351533386856e-05
step: 450, loss: 6.343043060041964e-05
step: 460, loss: 1.016995247482555e-05
step: 470, loss: 1.1473750419099815e-05
step: 480, loss: 0.00019667230662889779
step: 490, loss: 1.1689814527926501e-05
step: 500, loss: 1.2807156053895596e-05

```

step: 510, loss: 1.6323507225024514e-05
step: 520, loss: 3.448736242717132e-05
step: 530, loss: 5.705454168491997e-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=0.9480037140204272, f1=0.9417206290471785,
best_f1=0.9417206290471785
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, 253.12it/s]
load_f1 = 0.9436422915696321
real_f1 = 0.9465364946536494
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:21, 200.64it/s]

3.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5315355062484741
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.3499005138874054
step: 20, loss: 0.4252147972583771
step: 30, loss: 0.36735621094703674
step: 40, loss: 0.31996649503707886
step: 50, loss: 0.5051276087760925
step: 60, loss: 0.5092936754226685
step: 70, loss: 0.2943290174007416
step: 80, loss: 0.36444466996192932
step: 90, loss: 0.23546600341796875
step: 100, loss: 0.24102665483951569
step: 110, loss: 0.22962349653244019
step: 120, loss: 0.40549004077911377
step: 130, loss: 0.32080554962158203
step: 140, loss: 0.5054018497467041
step: 150, loss: 0.3685460090637207
step: 160, loss: 0.48921364545822144
step: 170, loss: 0.23853202164173126
step: 180, loss: 0.37568241357803345
step: 190, loss: 0.6147156953811646

```

```

step: 200, loss: 0.36843687295913696
step: 210, loss: 0.5107548236846924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38760364055633545
step: 10, loss: 0.18549123406410217
step: 20, loss: 0.47378307580947876
step: 30, loss: 0.4849238395690918
step: 40, loss: 0.5048852562904358
step: 50, loss: 0.22793850302696228
step: 60, loss: 0.31319501996040344
step: 70, loss: 0.4484672546386719
step: 80, loss: 0.3078884780406952
step: 90, loss: 0.3835938274860382
step: 100, loss: 0.5319280624389648
step: 110, loss: 0.3814009130001068
step: 120, loss: 0.25711357593536377
step: 130, loss: 0.15503616631031036
step: 140, loss: 0.23648394644260406
step: 150, loss: 0.4256758689880371
step: 160, loss: 0.1565236896276474
step: 170, loss: 0.6087914705276489
step: 180, loss: 0.3380194306373596
step: 190, loss: 0.3191121220588684
step: 200, loss: 0.15257106721401215
step: 210, loss: 0.3155454993247986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26087409257888794
step: 10, loss: 0.2424335926771164
step: 20, loss: 0.4745539724826813
step: 30, loss: 0.25502488017082214

```

```

step: 40, loss: 0.43818017840385437
step: 50, loss: 0.473752498626709
step: 60, loss: 0.4924561381340027
step: 70, loss: 0.18870893120765686
step: 80, loss: 0.4692935645580292
step: 90, loss: 0.2401108294725418
step: 100, loss: 0.3995683491230011
step: 110, loss: 0.2346152365207672
step: 120, loss: 0.2330617755651474
step: 130, loss: 0.1479753851890564
step: 140, loss: 0.4006103277206421
step: 150, loss: 0.323042631149292
step: 160, loss: 0.21625329554080963
step: 170, loss: 0.3794671893119812
step: 180, loss: 0.25840863585472107
step: 190, loss: 0.16835078597068787
step: 200, loss: 0.2374555617570877
step: 210, loss: 0.2553613781929016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3146429657936096
step: 10, loss: 0.31248530745506287
step: 20, loss: 0.30915960669517517
step: 30, loss: 0.2578415870666504
step: 40, loss: 0.23427173495292664
step: 50, loss: 0.24885664880275726
step: 60, loss: 0.4908374547958374
step: 70, loss: 0.2531527578830719
step: 80, loss: 0.23854181170463562
step: 90, loss: 0.4456881582736969
step: 100, loss: 0.37863531708717346
step: 110, loss: 0.6339535117149353
step: 120, loss: 0.37196606397628784
step: 130, loss: 0.6819242835044861
step: 140, loss: 0.5011682510375977
step: 150, loss: 0.37829822301864624
step: 160, loss: 0.31387391686439514
step: 170, loss: 0.16917301714420319
step: 180, loss: 0.08158579468727112
step: 190, loss: 0.16586625576019287
step: 200, loss: 0.3126537799835205
step: 210, loss: 0.37366950511932373

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38427478075027466
step: 10, loss: 0.32066938281059265
step: 20, loss: 0.309684157371521
step: 30, loss: 0.23823140561580658
step: 40, loss: 0.44515153765678406
step: 50, loss: 0.3808741569519043
step: 60, loss: 0.38220006227493286
step: 70, loss: 0.24363970756530762
step: 80, loss: 0.4357181787490845
step: 90, loss: 0.4466342329978943
step: 100, loss: 0.24217358231544495
step: 110, loss: 0.16899096965789795
step: 120, loss: 0.24227671325206757
step: 130, loss: 0.3115865886211395
step: 140, loss: 0.534923791885376
step: 150, loss: 0.3129124343395233
step: 160, loss: 0.24393950402736664
step: 170, loss: 0.377014696598053
step: 180, loss: 0.23467399179935455
step: 190, loss: 0.525456428527832
step: 200, loss: 0.38008758425712585
step: 210, loss: 0.24571967124938965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17152363061904907
step: 10, loss: 0.3801587224006653
step: 20, loss: 0.3089004456996918
step: 30, loss: 0.2424665242433548
step: 40, loss: 0.31274154782295227
step: 50, loss: 0.6031544208526611
step: 60, loss: 0.30770811438560486
step: 70, loss: 0.4364834725856781
step: 80, loss: 0.31466320157051086
step: 90, loss: 0.3135385513305664

```

```

step: 100, loss: 0.3184583783149719
step: 110, loss: 0.37929877638816833
step: 120, loss: 0.24813590943813324
step: 130, loss: 0.24020253121852875
step: 140, loss: 0.3806005120277405
step: 150, loss: 0.1894434094429016
step: 160, loss: 0.16667446494102478
step: 170, loss: 0.4580152630805969
step: 180, loss: 0.3847077488899231
step: 190, loss: 0.44839876890182495
step: 200, loss: 0.31833019852638245
step: 210, loss: 0.3730833828449249
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44985249638557434
step: 10, loss: 0.3180447816848755
step: 20, loss: 0.433015912771225
step: 30, loss: 0.24411670863628387
step: 40, loss: 0.2441760003566742
step: 50, loss: 0.3766874670982361
step: 60, loss: 0.3132193982601166
step: 70, loss: 0.24872459471225739
step: 80, loss: 0.4970589578151703
step: 90, loss: 0.3740490674972534
step: 100, loss: 0.45071497559547424
step: 110, loss: 0.30896586179733276
step: 120, loss: 0.308864027261734
step: 130, loss: 0.46796557307243347
step: 140, loss: 0.3160597085952759
step: 150, loss: 0.31170469522476196
step: 160, loss: 0.5365056991577148
step: 170, loss: 0.6290891170501709
step: 180, loss: 0.24377137422561646
step: 190, loss: 0.24446618556976318
step: 200, loss: 0.3087470829486847
step: 210, loss: 0.3142232596874237
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6188668608665466
step: 10, loss: 0.3146820664405823
step: 20, loss: 0.2542860805988312
step: 30, loss: 0.25917574763298035
step: 40, loss: 0.2392612248659134
step: 50, loss: 0.1658327281475067
step: 60, loss: 0.1644020825624466
step: 70, loss: 0.4489828646183014
step: 80, loss: 0.31162065267562866
step: 90, loss: 0.3931407332420349
step: 100, loss: 0.6067689657211304
step: 110, loss: 0.3709392845630646
step: 120, loss: 0.3074938654899597
step: 130, loss: 0.16736820340156555
step: 140, loss: 0.31386780738830566
step: 150, loss: 0.45419853925704956
step: 160, loss: 0.4600842297077179
step: 170, loss: 0.5249841809272766
step: 180, loss: 0.31350430846214294
step: 190, loss: 0.23724500834941864
step: 200, loss: 0.4636078178882599
step: 210, loss: 0.3788348436355591
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5684298872947693
step: 10, loss: 0.27286797761917114
step: 20, loss: 0.3760156035423279
step: 30, loss: 0.16810280084609985
step: 40, loss: 0.16734609007835388
step: 50, loss: 0.46289849281311035
step: 60, loss: 0.17290812730789185
step: 70, loss: 0.3793628215789795
step: 80, loss: 0.31536832451820374
step: 90, loss: 0.3129365146160126
step: 100, loss: 0.4718785881996155
step: 110, loss: 0.5221935510635376
step: 120, loss: 0.3718511164188385
step: 130, loss: 0.2402617186307907
step: 140, loss: 0.6055778861045837
step: 150, loss: 0.11527569591999054

```

```

step: 160, loss: 0.37332218885421753
step: 170, loss: 0.31218111515045166
step: 180, loss: 0.4611310362815857
step: 190, loss: 0.31495529413223267
step: 200, loss: 0.3094131350517273
step: 210, loss: 0.316491961479187
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3166620135307312
step: 10, loss: 0.3817645311355591
step: 20, loss: 0.25408124923706055
step: 30, loss: 0.17191879451274872
step: 40, loss: 0.30925098061561584
step: 50, loss: 0.5174348950386047
step: 60, loss: 0.24736164510250092
step: 70, loss: 0.3873487710952759
step: 80, loss: 0.16391238570213318
step: 90, loss: 0.393140584230423
step: 100, loss: 0.4623483419418335
step: 110, loss: 0.17573712766170502
step: 120, loss: 0.44752037525177
step: 130, loss: 0.23987431824207306
step: 140, loss: 0.3809804916381836
step: 150, loss: 0.24973411858081818
step: 160, loss: 0.3744177520275116
step: 170, loss: 0.24858316779136658
step: 180, loss: 0.3108537197113037
step: 190, loss: 0.24503876268863678
step: 200, loss: 0.6512852311134338
step: 210, loss: 0.24685943126678467
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3189600110054016
step: 10, loss: 0.30833813548088074
step: 20, loss: 0.38146138191223145
step: 30, loss: 0.1666651964187622

```

```

step: 40, loss: 0.44444334506988525
step: 50, loss: 0.44580668210983276
step: 60, loss: 0.43708527088165283
step: 70, loss: 0.1720099002122879
step: 80, loss: 0.5197120308876038
step: 90, loss: 0.38327139616012573
step: 100, loss: 0.5123281478881836
step: 110, loss: 0.4375613331794739
step: 120, loss: 0.42947086691856384
step: 130, loss: 0.24369299411773682
step: 140, loss: 0.3079323470592499
step: 150, loss: 0.309073269367218
step: 160, loss: 0.17096352577209473
step: 170, loss: 0.30232563614845276
step: 180, loss: 0.2368156611919403
step: 190, loss: 0.46313849091529846
step: 200, loss: 0.179498091340065
step: 210, loss: 0.384324312210083
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24802012741565704
step: 10, loss: 0.3124389052391052
step: 20, loss: 0.44860222935676575
step: 30, loss: 0.3129611015319824
step: 40, loss: 0.17430275678634644
step: 50, loss: 0.461868554353714
step: 60, loss: 0.10720124840736389
step: 70, loss: 0.4553547203540802
step: 80, loss: 0.31727173924446106
step: 90, loss: 0.3795640468597412
step: 100, loss: 0.12308503687381744
step: 110, loss: 0.24650873243808746
step: 120, loss: 0.24826280772686005
step: 130, loss: 0.44606828689575195
step: 140, loss: 0.38877952098846436
step: 150, loss: 0.09979187697172165
step: 160, loss: 0.2402307242155075
step: 170, loss: 0.23860464990139008
step: 180, loss: 0.30791589617729187
step: 190, loss: 0.37729960680007935
step: 200, loss: 0.18198207020759583
step: 210, loss: 0.311454713344574

```

```

/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30762147903442383
step: 10, loss: 0.24525025486946106
step: 20, loss: 0.31036579608917236
step: 30, loss: 0.2417009323835373
step: 40, loss: 0.23504488170146942
step: 50, loss: 0.24888929724693298
step: 60, loss: 0.5664266347885132
step: 70, loss: 0.4445507824420929
step: 80, loss: 0.38228368759155273
step: 90, loss: 0.24770338833332062
step: 100, loss: 0.3774382174015045
step: 110, loss: 0.24170038104057312
step: 120, loss: 0.31677985191345215
step: 130, loss: 0.3038482666015625
step: 140, loss: 0.15670232474803925
step: 150, loss: 0.31309831142425537
step: 160, loss: 0.5150550603866577
step: 170, loss: 0.22878000140190125
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 180, loss: 0.2192479521036148
step: 190, loss: 0.25315311551094055
step: 200, loss: 0.349228173494339
step: 210, loss: 0.38217616081237793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.358814352574103, f1=0.36296296296296293,
best_f1=0.36296296296296293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27447569370269775
step: 10, loss: 0.3556302487850189
step: 20, loss: 0.37422654032707214
step: 30, loss: 0.1693706512451172
step: 40, loss: 0.25987792015075684

```

```

step: 50, loss: 0.15674038231372833
step: 60, loss: 0.37710797786712646
step: 70, loss: 0.23650117218494415
step: 80, loss: 0.2611631453037262
step: 90, loss: 0.1467832326889038
step: 100, loss: 0.21379926800727844
step: 110, loss: 0.3040474057197571
step: 120, loss: 0.12469776719808578
step: 130, loss: 0.26312437653541565
step: 140, loss: 0.26231902837753296
step: 150, loss: 0.33992502093315125
step: 160, loss: 0.16302333772182465
step: 170, loss: 0.3210800290107727
step: 180, loss: 0.17181028425693512
step: 190, loss: 0.3488447368144989
step: 200, loss: 0.08858992159366608
step: 210, loss: 0.2542662024497986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4227642276422764, f1=0.4467455621301775,
best_f1=0.4467455621301775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.200435608625412
step: 10, loss: 0.22290873527526855
step: 20, loss: 0.368196040391922
step: 30, loss: 0.18119968473911285
step: 40, loss: 0.3476187586784363
step: 50, loss: 0.12682382762432098
step: 60, loss: 0.21839572489261627
step: 70, loss: 0.29023292660713196
step: 80, loss: 0.17228642106056213
step: 90, loss: 0.35137808322906494
step: 100, loss: 0.18472394347190857
step: 110, loss: 0.22547884285449982
step: 120, loss: 0.235802561044693
step: 130, loss: 0.1163415014743805
step: 140, loss: 0.2672272324562073
step: 150, loss: 0.36283639073371887
step: 160, loss: 0.3935573995113373
step: 170, loss: 0.2915768623352051
step: 180, loss: 0.30898311734199524

```

```

step: 190, loss: 0.22509966790676117
step: 200, loss: 0.2239765077829361
step: 210, loss: 0.37456199526786804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.43859649122807015, f1=0.4577861163227017,
best_f1=0.4577861163227017
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, 476.77it/s]
load_f1 = 0.4409171075837742
real_f1 = 0.42628774422735344
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:21, 203.18it/s]

3.1.5 Walmart-Amazon - Running the matcher

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

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

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

Selected optimization level 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: 0.45218509435653687
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.38263705372810364
step: 20, loss: 0.25064441561698914
step: 30, loss: 0.38666439056396484
step: 40, loss: 0.29327836632728577
step: 50, loss: 0.3173375725746155
step: 60, loss: 0.4395466446876526
step: 70, loss: 0.40926679968833923
step: 80, loss: 0.15449059009552002
step: 90, loss: 0.3336421549320221
step: 100, loss: 0.43028929829597473
step: 110, loss: 0.2397184520959854
step: 120, loss: 0.33727943897247314
step: 130, loss: 0.33071163296699524
step: 140, loss: 0.17822693288326263
step: 150, loss: 0.2965547442436218
step: 160, loss: 0.24453139305114746
step: 170, loss: 0.38730698823928833
step: 180, loss: 0.15317818522453308
step: 190, loss: 0.15858513116836548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3839445114135742
step: 10, loss: 0.32515230774879456
step: 20, loss: 0.6330191493034363
step: 30, loss: 0.24513089656829834
step: 40, loss: 0.5624115467071533
step: 50, loss: 0.3128304183483124
step: 60, loss: 0.4387086033821106
step: 70, loss: 0.38141852617263794
step: 80, loss: 0.14201907813549042
step: 90, loss: 0.3088798224925995
step: 100, loss: 0.2487298846244812
step: 110, loss: 0.3320527672767639
step: 120, loss: 0.22354434430599213
step: 130, loss: 0.43380406498908997
step: 140, loss: 0.35862359404563904
step: 150, loss: 0.31494826078414917
step: 160, loss: 0.28748995065689087
step: 170, loss: 0.24028639495372772
step: 180, loss: 0.1724352389574051
step: 190, loss: 0.22467657923698425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20195439739413681, f1=0.20568927789934355,
best_f1=0.20568927789934355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36848413944244385
step: 10, loss: 0.35296404361724854
step: 20, loss: 0.42134934663772583
step: 30, loss: 0.29828616976737976
step: 40, loss: 0.09994648396968842

```

```

step: 50, loss: 0.3642648160457611
step: 60, loss: 0.1541454941034317
step: 70, loss: 0.38041192293167114
step: 80, loss: 0.28537872433662415
step: 90, loss: 0.37490203976631165
step: 100, loss: 0.5210122466087341
step: 110, loss: 0.6079935431480408
step: 120, loss: 0.36225762963294983
step: 130, loss: 0.22296860814094543
step: 140, loss: 0.284537672996521
step: 150, loss: 0.40674781799316406
step: 160, loss: 0.1922273188829422
step: 170, loss: 0.4152910113334656
step: 180, loss: 0.32430610060691833
step: 190, loss: 0.18878808617591858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5804988662131519, f1=0.5773672055427252,
best_f1=0.5773672055427252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11112629622220993
step: 10, loss: 0.1899355947971344
step: 20, loss: 0.3604242205619812
step: 30, loss: 0.22711805999279022
step: 40, loss: 0.47994565963745117
step: 50, loss: 0.22887596487998962
step: 60, loss: 0.33903804421424866
step: 70, loss: 0.27910447120666504
step: 80, loss: 0.20422489941120148
step: 90, loss: 0.07741469889879227
step: 100, loss: 0.20776158571243286
step: 110, loss: 0.30294162034988403
step: 120, loss: 0.05445515364408493
step: 130, loss: 0.18930530548095703
step: 140, loss: 0.3521345257759094
step: 150, loss: 0.09061479568481445
step: 160, loss: 0.1930999606847763
step: 170, loss: 0.16822850704193115
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 180, loss: 0.13117392361164093
step: 190, loss: 0.08333764225244522

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.678132678132678, f1=0.6486486486486486,
best_f1=0.6486486486486486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23261387646198273
step: 10, loss: 0.216104194521904
step: 20, loss: 0.0778663232922554
step: 30, loss: 0.26783713698387146
step: 40, loss: 0.12523245811462402
step: 50, loss: 0.11698152869939804
step: 60, loss: 0.02264256589114666
step: 70, loss: 0.10916295647621155
step: 80, loss: 0.14548444747924805
step: 90, loss: 0.15256479382514954
step: 100, loss: 0.21249458193778992
step: 110, loss: 0.14088939130306244
step: 120, loss: 0.12023565918207169
step: 130, loss: 0.40021204948425293
step: 140, loss: 0.07936512678861618
step: 150, loss: 0.19128145277500153
step: 160, loss: 0.04736102744936943
step: 170, loss: 0.06620191782712936
step: 180, loss: 0.05369015783071518
step: 190, loss: 0.10480596870183945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.711484593837535, f1=0.718918918918919,
best_f1=0.718918918918919
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14449505507946014
step: 10, loss: 0.18091730773448944
step: 20, loss: 0.17196525633335114
step: 30, loss: 0.2039417028427124

```

```

step: 40, loss: 0.06974392384290695
step: 50, loss: 0.06994021683931351
step: 60, loss: 0.11285365372896194
step: 70, loss: 0.15980826318264008
step: 80, loss: 0.19361189007759094
step: 90, loss: 0.10079428553581238
step: 100, loss: 0.09084479510784149
step: 110, loss: 0.051096267998218536
step: 120, loss: 0.08053570985794067
step: 130, loss: 0.20982573926448822
step: 140, loss: 0.11357560753822327
step: 150, loss: 0.07260176539421082
step: 160, loss: 0.15907517075538635
step: 170, loss: 0.33161434531211853
step: 180, loss: 0.042028263211250305
step: 190, loss: 0.010173220187425613
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7874015748031497, f1=0.7641025641025642,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1954323798418045
step: 10, loss: 0.015172833576798439
step: 20, loss: 0.04512249678373337
step: 30, loss: 0.024896329268813133
step: 40, loss: 0.056843943893909454
step: 50, loss: 0.01338243205100298
step: 60, loss: 0.17858755588531494
step: 70, loss: 0.012075331062078476
step: 80, loss: 0.008033854886889458
step: 90, loss: 0.053669802844524384
step: 100, loss: 0.2909972369670868
step: 110, loss: 0.3393624722957611
step: 120, loss: 0.18935418128967285
step: 130, loss: 0.050958242267370224
step: 140, loss: 0.09178269654512405
step: 150, loss: 0.024860480800271034
step: 160, loss: 0.4205145537853241
step: 170, loss: 0.05260228365659714
step: 180, loss: 0.05107530206441879
step: 190, loss: 0.1561741679906845

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7890410958904108, f1=0.7798408488063661,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047755077481269836
step: 10, loss: 0.08203796297311783
step: 20, loss: 0.006669482681900263
step: 30, loss: 0.1789967566728592
step: 40, loss: 0.08637892454862595
step: 50, loss: 0.02021612413227558
step: 60, loss: 0.11582549661397934
step: 70, loss: 0.06713084876537323
step: 80, loss: 0.21349596977233887
step: 90, loss: 0.021064404398202896
step: 100, loss: 0.033217813819646835
step: 110, loss: 0.010470491833984852
step: 120, loss: 0.007645849138498306
step: 130, loss: 0.010758318938314915
step: 140, loss: 0.012730980291962624
step: 150, loss: 0.06793061643838882
step: 160, loss: 0.002860400127246976
step: 170, loss: 0.002961871912702918
step: 180, loss: 0.1007644459605217
step: 190, loss: 0.020127492025494576
/usr/local/lib/python3.7/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=0.7592067988668556, f1=0.7722222222222223,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018556440249085426
step: 10, loss: 0.0019049065886065364
step: 20, loss: 0.017027825117111206
step: 30, loss: 0.004367087502032518
step: 40, loss: 0.13216672837734222
step: 50, loss: 0.018629079684615135
step: 60, loss: 0.03247866779565811
step: 70, loss: 0.0016312748193740845

```

```

step: 80, loss: 0.06774712353944778
step: 90, loss: 0.09257494658231735
step: 100, loss: 0.05945245549082756
step: 110, loss: 0.0024997596628963947
step: 120, loss: 0.023844296112656593
step: 130, loss: 0.03543291985988617
step: 140, loss: 0.07768528908491135
step: 150, loss: 0.015458088368177414
step: 160, loss: 0.009795461781322956
step: 170, loss: 0.11584419757127762
step: 180, loss: 0.1290440410375595
step: 190, loss: 0.0008066431619226933
/usr/local/lib/python3.7/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=0.7688022284122563, f1=0.7704918032786885,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026405546814203262
step: 10, loss: 0.08100218325853348
step: 20, loss: 0.02117081545293331
step: 30, loss: 0.1580880731344223
step: 40, loss: 0.004118049517273903
step: 50, loss: 0.004587765783071518
step: 60, loss: 0.012232924811542034
step: 70, loss: 0.0012216244358569384
step: 80, loss: 0.09071569889783859
step: 90, loss: 0.11957307159900665
step: 100, loss: 0.013748623430728912
step: 110, loss: 0.015531152486801147
step: 120, loss: 0.007202194072306156
step: 130, loss: 0.014567971229553223
step: 140, loss: 0.007160239387303591
step: 150, loss: 0.0010941321961581707
step: 160, loss: 0.0012356977676972747
step: 170, loss: 0.0011697810841724277
step: 180, loss: 0.010174596682190895
step: 190, loss: 0.0013569752918556333
/usr/local/lib/python3.7/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=0.751842751842752, f1=0.7903614457831326,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.007581330835819244
step: 10, loss: 0.000706033140886575
step: 20, loss: 0.02177845686674118
step: 30, loss: 0.10729959607124329
step: 40, loss: 0.009513665921986103
step: 50, loss: 0.0790453776717186
step: 60, loss: 0.027048494666814804
step: 70, loss: 0.007664382923394442
step: 80, loss: 0.024912046268582344
step: 90, loss: 0.0007218524697236717
step: 100, loss: 0.012500764802098274
step: 110, loss: 0.22319623827934265
step: 120, loss: 0.003083036281168461
step: 130, loss: 0.01671745255589485
step: 140, loss: 0.0014616405824199319
step: 150, loss: 0.0006209634593687952
step: 160, loss: 0.0070282211527228355
step: 170, loss: 0.016097070649266243
step: 180, loss: 0.002411956200376153
step: 190, loss: 0.0012615055311471224
/usr/local/lib/python3.7/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=0.7572815533980581, f1=0.7855421686746988,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009317221119999886
step: 10, loss: 0.002051897579804063
step: 20, loss: 0.001872443244792521
step: 30, loss: 0.06406436115503311
step: 40, loss: 0.0025503423530608416
step: 50, loss: 0.0021030178759247065
step: 60, loss: 0.003540209960192442
step: 70, loss: 0.0019494533771649003
step: 80, loss: 0.004765939898788929
step: 90, loss: 0.029665259644389153
step: 100, loss: 0.036150332540273666
step: 110, loss: 0.002300878055393696
step: 120, loss: 0.026528168469667435
step: 130, loss: 0.0009697844507172704
step: 140, loss: 0.002732753986492753
step: 150, loss: 0.029109662398695946
step: 160, loss: 0.0003046217898372561
step: 170, loss: 0.090687595307827
step: 180, loss: 0.002484117401763797
step: 190, loss: 0.0004948268760927022

```

```

/usr/local/lib/python3.7/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=0.7624020887728459, f1=0.7936507936507936,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009087341837584972
step: 10, loss: 0.014104277826845646
step: 20, loss: 0.004029245115816593
step: 30, loss: 0.05621296539902687
step: 40, loss: 0.0007821600302122533
step: 50, loss: 0.036575041711330414
step: 60, loss: 0.00083127774996683
step: 70, loss: 0.2202511578798294
step: 80, loss: 0.00422333087772131
step: 90, loss: 0.003566104918718338
step: 100, loss: 0.003132594982162118
step: 110, loss: 0.0009382432326674461
step: 120, loss: 0.007617168594151735
step: 130, loss: 0.002326982794329524
step: 140, loss: 0.024920281022787094
step: 150, loss: 0.0015794303035363555
step: 160, loss: 0.0263321865350008
step: 170, loss: 0.014670399017632008
step: 180, loss: 0.0008866036660037935
step: 190, loss: 0.18796151876449585
/usr/local/lib/python3.7/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=0.7708894878706198, f1=0.8238482384823848,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005645402241498232
step: 10, loss: 0.0017356242751702666
step: 20, loss: 0.012232023291289806
step: 30, loss: 0.014299317263066769
step: 40, loss: 0.02784414030611515
step: 50, loss: 0.025508670136332512
step: 60, loss: 0.0011230978416278958
step: 70, loss: 0.0006972087430767715
step: 80, loss: 0.01293440442532301
step: 90, loss: 0.0003865486360155046
step: 100, loss: 0.0009446366457268596
step: 110, loss: 0.0004177768714725971

```



```

step: 120, loss: 0.0025249351747334003
step: 130, loss: 0.0005407819990068674
step: 140, loss: 0.0006905944901518524
step: 150, loss: 0.001989776501432061
step: 160, loss: 0.001690203556790948
step: 170, loss: 0.0005046672886237502
step: 180, loss: 0.003980179782956839
step: 190, loss: 0.013517669402062893
/usr/local/lib/python3.7/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=0.7651715039577835, f1=0.8172043010752689,
best_f1=0.7798408488063661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0691433995962143
step: 10, loss: 0.0007967408164404333
step: 20, loss: 0.014664246700704098
step: 30, loss: 0.00038303693872876465
step: 40, loss: 0.0024587290827184916
step: 50, loss: 0.009022059850394726
step: 60, loss: 0.0002613542601466179
step: 70, loss: 0.0027942766901105642
step: 80, loss: 0.0003940882452297956
step: 90, loss: 0.004794737324118614
step: 100, loss: 0.0006277754437178373
step: 110, loss: 0.00413987971842289
step: 120, loss: 0.0010414196876809
step: 130, loss: 0.0006137517630122602
step: 140, loss: 0.0003480548330117017
step: 150, loss: 0.0002945430460385978
step: 160, loss: 0.00033259770134463906
step: 170, loss: 0.0007483543013222516
step: 180, loss: 0.00046928858500905335
step: 190, loss: 0.005029631312936544
/usr/local/lib/python3.7/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=0.7789473684210525, f1=0.8279569892473119,
best_f1=0.7798408488063661
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, 237.79it/s]
load_f1 = 0.8091168091168092
real_f1 = 0.8089887640449438
733it [00:00, 3361.97it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 195.39it/s]

```

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

```

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

```

```

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

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

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

```

Selected optimization level 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: 0.49937504529953003

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

step: 10, loss: 0.4275150001049042

step: 20, loss: 0.40234410762786865

step: 30, loss: 0.4756385385990143

step: 40, loss: 0.5641525983810425

step: 50, loss: 0.3036254346370697

step: 60, loss: 0.6184852719306946

step: 70, loss: 0.31947317719459534

step: 80, loss: 0.30438506603240967

step: 90, loss: 0.24379576742649078

step: 100, loss: 0.14649933576583862

step: 110, loss: 0.41008925437927246

step: 120, loss: 0.3187376856803894

step: 130, loss: 0.3059491515159607

step: 140, loss: 0.37817683815956116

step: 150, loss: 0.34612953662872314

step: 160, loss: 0.42336714267730713

step: 170, loss: 0.3174005448818207

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

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

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

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=0.2603036876355749, f1=0.20952380952380953,

best_f1=0.20952380952380953

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

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

step: 0, loss: 0.2929944097995758

step: 10, loss: 0.43367934226989746

step: 20, loss: 0.3130885660648346

step: 30, loss: 0.2905385494232178

step: 40, loss: 0.06434380263090134

step: 50, loss: 0.3697313368320465

step: 60, loss: 0.1819867491722107

```

step: 70, loss: 0.45523953437805176
step: 80, loss: 0.20762793719768524
step: 90, loss: 0.1884293407201767
step: 100, loss: 0.2521023750305176
step: 110, loss: 0.2925751805305481
step: 120, loss: 0.14343540370464325
step: 130, loss: 0.2318824976682663
step: 140, loss: 0.3601487874984741
step: 150, loss: 0.15548524260520935
step: 160, loss: 0.37842124700546265
step: 170, loss: 0.2958607077598572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6649874055415618, f1=0.7066974595842955,
best_f1=0.7066974595842955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3725753724575043
step: 10, loss: 0.0998145341873169
step: 20, loss: 0.07271026074886322
step: 30, loss: 0.10694018006324768
step: 40, loss: 0.22157524526119232
step: 50, loss: 0.30784469842910767
step: 60, loss: 0.07565142214298248
step: 70, loss: 0.12010953575372696
step: 80, loss: 0.11204785853624344
step: 90, loss: 0.4098859131336212
step: 100, loss: 0.08519889414310455
step: 110, loss: 0.13344520330429077
step: 120, loss: 0.1912878155708313
step: 130, loss: 0.10177198052406311
step: 140, loss: 0.17943917214870453
step: 150, loss: 0.07849371433258057
step: 160, loss: 0.20834405720233917
step: 170, loss: 0.07514951378107071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7715736040609138, f1=0.8232445520581113,
best_f1=0.8232445520581113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2386096715927124
step: 10, loss: 0.22627249360084534
step: 20, loss: 0.012266858480870724
step: 30, loss: 0.4139823019504547
step: 40, loss: 0.1034686490893364
step: 50, loss: 0.1658208668231964
step: 60, loss: 0.2540101110935211
step: 70, loss: 0.011504010297358036
step: 80, loss: 0.213097482919693
step: 90, loss: 0.17152154445648193
step: 100, loss: 0.06924621015787125
step: 110, loss: 0.22891327738761902
step: 120, loss: 0.2553721070289612
step: 130, loss: 0.07496441155672073
step: 140, loss: 0.15629054605960846
step: 150, loss: 0.24092963337898254
step: 160, loss: 0.03970453888177872
step: 170, loss: 0.02757698856294155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7772511848341233, f1=0.8125000000000001,
best_f1=0.8125000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015000542625784874
step: 10, loss: 0.021727904677391052
step: 20, loss: 0.057015687227249146
step: 30, loss: 0.08918452262878418
step: 40, loss: 0.07221991568803787
step: 50, loss: 0.03063199296593666
step: 60, loss: 0.0469801090657711
step: 70, loss: 0.03679788112640381
step: 80, loss: 0.004017567727714777
step: 90, loss: 0.019980797544121742
step: 100, loss: 0.03153756633400917
step: 110, loss: 0.20673546195030212
step: 120, loss: 0.03387269750237465

```

```

step: 130, loss: 0.003623182652518153
step: 140, loss: 0.06820692867040634
step: 150, loss: 0.1532878577709198
step: 160, loss: 0.03378194570541382
step: 170, loss: 0.04198151454329491
/usr/local/lib/python3.7/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=0.7684964200477328, f1=0.8381374722838137,
best_f1=0.8125000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23122826218605042
step: 10, loss: 0.03542110323905945
step: 20, loss: 0.11163505166769028
step: 30, loss: 0.011497801169753075
step: 40, loss: 0.006789067760109901
step: 50, loss: 0.006014656741172075
step: 60, loss: 0.06702740490436554
step: 70, loss: 0.028115754947066307
step: 80, loss: 0.06979053467512131
step: 90, loss: 0.15835627913475037
step: 100, loss: 0.07115334272384644
step: 110, loss: 0.20926353335380554
step: 120, loss: 0.029500780627131462
step: 130, loss: 0.13553500175476074
step: 140, loss: 0.2920927405357361
step: 150, loss: 0.05046885460615158
step: 160, loss: 0.07731255143880844
step: 170, loss: 0.11216545104980469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8049382716049382, f1=0.8459770114942529,
best_f1=0.8459770114942529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023455381393432617
step: 10, loss: 0.024634432047605515
step: 20, loss: 0.0063246022909879684
step: 30, loss: 0.005114036612212658
step: 40, loss: 0.007274768315255642

```

```

step: 50, loss: 0.005602596327662468
step: 60, loss: 0.13197101652622223
step: 70, loss: 0.00953297782689333
step: 80, loss: 0.025090858340263367
step: 90, loss: 0.16304321587085724
step: 100, loss: 0.00722188176587224
step: 110, loss: 0.08911830186843872
step: 120, loss: 0.011157797649502754
step: 130, loss: 0.11005125194787979
step: 140, loss: 0.03364790976047516
step: 150, loss: 0.1332588791847229
step: 160, loss: 0.010216000489890575
step: 170, loss: 0.08568821847438812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8244274809160305, f1=0.8605200945626478,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10028894245624542
step: 10, loss: 0.09045973420143127
step: 20, loss: 0.012931648641824722
step: 30, loss: 0.0018078357679769397
step: 40, loss: 0.014848195016384125
step: 50, loss: 0.007687960285693407
step: 60, loss: 0.010322131216526031
step: 70, loss: 0.023044729605317116
step: 80, loss: 0.042259931564331055
step: 90, loss: 0.1290188729763031
step: 100, loss: 0.009515068493783474
step: 110, loss: 0.14992959797382355
step: 120, loss: 0.048136476427316666
step: 130, loss: 0.003843864193186164
step: 140, loss: 0.03344237059354782
step: 150, loss: 0.0015204009832814336
step: 160, loss: 0.0016676485538482666
step: 170, loss: 0.05867405980825424
/usr/local/lib/python3.7/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=0.8019559902200489, f1=0.8425925925925926,
best_f1=0.8605200945626478

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036786384880542755
step: 10, loss: 0.010011817328631878
step: 20, loss: 0.005488230846822262
step: 30, loss: 0.010758132673799992
step: 40, loss: 0.20362339913845062
step: 50, loss: 0.0017019379884004593
step: 60, loss: 0.054128292948007584
step: 70, loss: 0.01902049221098423
step: 80, loss: 0.00173981545958668
step: 90, loss: 0.11433196812868118
step: 100, loss: 0.11999095976352692
step: 110, loss: 0.032828811556100845
step: 120, loss: 0.05807310342788696
step: 130, loss: 0.003437630832195282
step: 140, loss: 0.02405451238155365
step: 150, loss: 0.18835382163524628
step: 160, loss: 0.001967397751286626
step: 170, loss: 0.05239766836166382
/usr/local/lib/python3.7/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=0.7840375586854459, f1=0.800865800865801,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006945963017642498
step: 10, loss: 0.08502734452486038
step: 20, loss: 0.03720592334866524
step: 30, loss: 0.03287457674741745
step: 40, loss: 0.006865772884339094
step: 50, loss: 0.022027669474482536
step: 60, loss: 0.06971138715744019
step: 70, loss: 0.0009321744437329471
step: 80, loss: 0.029421521350741386
step: 90, loss: 0.0030785463750362396
step: 100, loss: 0.0012810752959921956
step: 110, loss: 0.007965335622429848
step: 120, loss: 0.0024230265989899635
step: 130, loss: 0.01716230809688568
step: 140, loss: 0.007461973465979099
step: 150, loss: 0.0025264353025704622
step: 160, loss: 0.0017675241688266397
step: 170, loss: 0.03772571682929993
/usr/local/lib/python3.7/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=0.8078817733990147, f1=0.828054298642534,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04420977458357811
step: 10, loss: 0.002571354154497385
step: 20, loss: 0.0014253529952839017
step: 30, loss: 0.021243268623948097
step: 40, loss: 0.001215803436934948
step: 50, loss: 0.05601734668016434
step: 60, loss: 0.1454978585243225
step: 70, loss: 0.010356487706303596
step: 80, loss: 0.01113049779087305
step: 90, loss: 0.005077308043837547
step: 100, loss: 0.02535357140004635
step: 110, loss: 0.006152545567601919
step: 120, loss: 0.003559491131454706
step: 130, loss: 0.00896878819912672
step: 140, loss: 0.00447113998234272
step: 150, loss: 0.0009983314666897058
step: 160, loss: 0.06840734928846359
step: 170, loss: 0.007946785539388657
/usr/local/lib/python3.7/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=0.810126582278481, f1=0.850467289719626,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007502064574509859
step: 10, loss: 0.0009862167062237859
step: 20, loss: 0.0011714197462424636
step: 30, loss: 0.02335827238857746
step: 40, loss: 0.0013518587220460176
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.005557131487876177
step: 60, loss: 0.04314843937754631
step: 70, loss: 0.0014728270471096039
step: 80, loss: 0.0010307321790605783
step: 90, loss: 0.017829839140176773
step: 100, loss: 0.015767131000757217
step: 110, loss: 0.007257724180817604
step: 120, loss: 0.04681169241666794
step: 130, loss: 0.0009858699049800634

```

```

step: 140, loss: 0.0008294585277326405
step: 150, loss: 0.0031469136010855436
step: 160, loss: 0.12181335687637329
step: 170, loss: 0.25644928216934204
/usr/local/lib/python3.7/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=0.7969151670951157, f1=0.8510638297872338,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014600445283576846
step: 10, loss: 0.000867595721501857
step: 20, loss: 0.0040026260539889336
step: 30, loss: 0.020600304007530212
step: 40, loss: 0.0012950679520145059
step: 50, loss: 0.048472799360752106
step: 60, loss: 0.0022251431364566088
step: 70, loss: 0.05537419021129608
step: 80, loss: 0.0004773248510900885
step: 90, loss: 0.010533139109611511
step: 100, loss: 0.011803270317614079
step: 110, loss: 0.004110512789338827
step: 120, loss: 0.010457768104970455
step: 130, loss: 0.002015067031607032
step: 140, loss: 0.029204079881310463
step: 150, loss: 0.0006738969241268933
step: 160, loss: 0.05306893214583397
step: 170, loss: 0.001878997660242021
/usr/local/lib/python3.7/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=0.803921568627451, f1=0.8433179723502305,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02629600651562214
step: 10, loss: 0.0006731961038894951
step: 20, loss: 0.004016890190541744
step: 30, loss: 0.0014681637985631824
step: 40, loss: 0.0006891353987157345
step: 50, loss: 0.0012326951837167144
step: 60, loss: 0.0006907220231369138
step: 70, loss: 0.0016460544429719448
step: 80, loss: 0.0008975822129286826
step: 90, loss: 0.0025646307040005922

```

```

step: 100, loss: 0.01959124580025673
step: 110, loss: 0.012770448811352253
step: 120, loss: 0.0009284888510592282
step: 130, loss: 0.019030701369047165
step: 140, loss: 0.026416705921292305
step: 150, loss: 0.013523164205253124
step: 160, loss: 0.0032631533686071634
step: 170, loss: 0.020745055750012398
/usr/local/lib/python3.7/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=0.789838337182448, f1=0.8214285714285715,
best_f1=0.8605200945626478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011674300767481327
step: 10, loss: 0.006118605844676495
step: 20, loss: 0.0009794814977794886
step: 30, loss: 0.0020243350882083178
step: 40, loss: 0.0038926093839108944
step: 50, loss: 0.0005282158381305635
step: 60, loss: 0.0009039915166795254
step: 70, loss: 0.025995612144470215
step: 80, loss: 0.004762364085763693
step: 90, loss: 0.0006864148890599608
step: 100, loss: 0.014006364159286022
step: 110, loss: 0.004178865812718868
step: 120, loss: 0.0007085077813826501
step: 130, loss: 0.001726232934743166
step: 140, loss: 0.01420100312680006
step: 150, loss: 0.0011407765559852123
step: 160, loss: 0.0007576916250400245
step: 170, loss: 0.0025571032892912626
/usr/local/lib/python3.7/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=0.7848699763593381, f1=0.8206278026905829,
best_f1=0.8605200945626478
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : O2
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:06, 315.74it/s]
load_f1 = 0.6590389016018307
real_f1 = 0.62882096069869
0it [00:00, ?it/s]/usr/local/lib/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, 188.45it/s]

```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.5983002185821533

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

```

step: 10, loss: 0.4190135896205902
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.5068652033805847
step: 30, loss: 0.3847917318344116
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.3712872564792633
step: 50, loss: 0.6820937395095825
step: 60, loss: 0.5353055000305176
step: 70, loss: 0.4598830044269562
step: 80, loss: 0.5875711441040039
step: 90, loss: 0.4487157464027405
step: 100, loss: 0.37896716594696045
step: 110, loss: 0.2915041446685791
step: 120, loss: 0.17685867846012115
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 130, loss: 0.10225196182727814
step: 140, loss: 0.0897841826081276
step: 150, loss: 0.14802975952625275
step: 160, loss: 0.015839887782931328
step: 170, loss: 0.3196024000644684
step: 180, loss: 0.21735401451587677
step: 190, loss: 0.5844372510910034
step: 200, loss: 0.04052238538861275
step: 210, loss: 0.06232898682355881
step: 220, loss: 0.09466299414634705
step: 230, loss: 0.014864863827824593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9294871794871795, f1=0.9405286343612335,
best_f1=0.9405286343612335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012191430665552616
step: 10, loss: 0.19842109084129333
step: 20, loss: 0.09902332723140717
step: 30, loss: 0.06715508550405502
step: 40, loss: 0.2925213575363159
step: 50, loss: 0.037752583622932434
step: 60, loss: 0.017529787495732307
step: 70, loss: 0.05625211074948311
step: 80, loss: 0.006025425158441067
step: 90, loss: 0.03870997205376625

```

```

step: 100, loss: 0.005436161067336798
step: 110, loss: 0.07309173047542572
step: 120, loss: 0.05381219461560249
step: 130, loss: 0.02314545027911663
step: 140, loss: 0.007689740974456072
step: 150, loss: 0.0838407650589943
step: 160, loss: 0.15255989134311676
step: 170, loss: 0.0056470842100679874
step: 180, loss: 0.19449593126773834
step: 190, loss: 0.07804638147354126
step: 200, loss: 0.09372663497924805
step: 210, loss: 0.0837504044175148
step: 220, loss: 0.006065691355615854
step: 230, loss: 0.003151955548673868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9568106312292359, f1=0.9530201342281879,
best_f1=0.9530201342281879
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0233306847512722
step: 10, loss: 0.04719466343522072
step: 20, loss: 0.03784126043319702
step: 30, loss: 0.011749106459319592
step: 40, loss: 0.01765735074877739
step: 50, loss: 0.004405859857797623
step: 60, loss: 0.15072232484817505
step: 70, loss: 0.03548286855220795
step: 80, loss: 0.05642553046345711
step: 90, loss: 0.022539988160133362
step: 100, loss: 0.06654862314462662
step: 110, loss: 0.0154570322483778
step: 120, loss: 0.0020431377924978733
step: 130, loss: 0.1266525834798813
step: 140, loss: 0.003975964151322842
step: 150, loss: 0.06844515353441238
step: 160, loss: 0.04799412563443184
step: 170, loss: 0.0011869887821376324
step: 180, loss: 0.008524131961166859
step: 190, loss: 0.04933521896600723
step: 200, loss: 0.013960499316453934
step: 210, loss: 0.003261391306295991

```



```

step: 220, loss: 0.07334921509027481
step: 230, loss: 0.012221897020936012
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9552572706935123,
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: 0.019192859530448914
step: 10, loss: 0.0014357790350914001
step: 20, loss: 0.02343621663749218
step: 30, loss: 0.018611015751957893
step: 40, loss: 0.02190844528377056
step: 50, loss: 0.07278323173522949
step: 60, loss: 0.08626578003168106
step: 70, loss: 0.03386720269918442
step: 80, loss: 0.19259345531463623
step: 90, loss: 0.026362745091319084
step: 100, loss: 0.029769547283649445
step: 110, loss: 0.004413907416164875
step: 120, loss: 0.03908532112836838
step: 130, loss: 0.008449846878647804
step: 140, loss: 0.005022221244871616
step: 150, loss: 0.010441645048558712
step: 160, loss: 0.008060193620622158
step: 170, loss: 0.007684089709073305
step: 180, loss: 0.029886411502957344
step: 190, loss: 0.004740601871162653
step: 200, loss: 0.0584050677716732
step: 210, loss: 0.032607171684503555
step: 220, loss: 0.0022930335253477097
step: 230, loss: 0.005171655211597681
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9647577092511014, f1=0.9559471365638766,
best_f1=0.9559471365638766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006402699509635568
step: 10, loss: 0.00843836646527052
step: 20, loss: 0.08506911993026733
step: 30, loss: 0.003286969382315874
step: 40, loss: 0.003632507985457778
step: 50, loss: 0.01377866044640541
step: 60, loss: 0.07463432103395462
step: 70, loss: 0.004280184395611286
step: 80, loss: 0.09453446418046951
step: 90, loss: 0.1630130559206009
step: 100, loss: 0.009518951177597046
step: 110, loss: 0.016838304698467255
step: 120, loss: 0.011086641810834408
step: 130, loss: 0.0123825678601861
step: 140, loss: 0.05067390576004982
step: 150, loss: 0.027252135798335075
step: 160, loss: 0.00646588858217001
step: 170, loss: 0.002323913387954235
step: 180, loss: 0.025546006858348846
step: 190, loss: 0.07976680994033813
step: 200, loss: 0.04385288059711456
step: 210, loss: 0.0026098978705704212
step: 220, loss: 0.00609110901132226
step: 230, loss: 0.0407852828502655
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9719416386083053, f1=0.9697648376259798,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009130281978286803
step: 10, loss: 0.004053996875882149
step: 20, loss: 0.001174023374915123
step: 30, loss: 0.0019501816714182496
step: 40, loss: 0.0002809154975693673
step: 50, loss: 0.0007141502574086189
step: 60, loss: 0.002215605927631259
step: 70, loss: 0.03501596674323082
step: 80, loss: 0.0015097426949068904
step: 90, loss: 0.006379920057952404

```

```

step: 100, loss: 0.012952180579304695
step: 110, loss: 0.018557878211140633
step: 120, loss: 0.000679278455208987
step: 130, loss: 0.0018055853433907032
step: 140, loss: 0.0012639588676393032
step: 150, loss: 0.001552005298435688
step: 160, loss: 0.06735093146562576
step: 170, loss: 0.0006008087657392025
step: 180, loss: 0.0025192389730364084
step: 190, loss: 0.19152843952178955
step: 200, loss: 0.01147734560072422
step: 210, loss: 0.008374971337616444
step: 220, loss: 0.0012143307831138372
step: 230, loss: 0.0011880352394655347
/usr/local/lib/python3.7/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=0.9698324022346367, f1=0.9585666293393057,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015007680049166083
step: 10, loss: 0.0006504050688818097
step: 20, loss: 0.0005718086613342166
step: 30, loss: 0.00020886397396679968
step: 40, loss: 0.000714725349098444
step: 50, loss: 0.0024770412128418684
step: 60, loss: 0.0005783154629170895
step: 70, loss: 0.0004967845161445439
step: 80, loss: 0.006803854834288359
step: 90, loss: 0.00037511842674575746
step: 100, loss: 0.0002826632698997855
step: 110, loss: 0.0009609807748347521
step: 120, loss: 0.0006920574232935905
step: 130, loss: 0.15572147071361542
step: 140, loss: 0.01131976768374443
step: 150, loss: 0.05216718837618828
step: 160, loss: 0.0004557759966701269
step: 170, loss: 0.008764060214161873
step: 180, loss: 0.0004459570627659559
step: 190, loss: 0.0009027415653690696
step: 200, loss: 0.014095190912485123
step: 210, loss: 0.0014711589319631457
step: 220, loss: 0.039100755006074905
step: 230, loss: 0.0015884099993854761
/usr/local/lib/python3.7/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=0.9658213891951488, f1=0.9621380846325166,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000373130344087258
step: 10, loss: 0.0014413099270313978
step: 20, loss: 0.006057901307940483
step: 30, loss: 0.0004969738074578345
step: 40, loss: 0.000429962674388662
step: 50, loss: 0.0005200236919336021
step: 60, loss: 0.0009234744939021766
step: 70, loss: 0.0008238460868597031
step: 80, loss: 0.017260540276765823
step: 90, loss: 0.003534185467287898
step: 100, loss: 0.00444470951333642
step: 110, loss: 0.021168498322367668
step: 120, loss: 0.009340728633105755
step: 130, loss: 0.003410627134144306
step: 140, loss: 0.001212811446748674
step: 150, loss: 0.07722315192222595
step: 160, loss: 0.00758645823225379
step: 170, loss: 0.0010714936070144176
step: 180, loss: 0.00029490291490219533
step: 190, loss: 0.002092963084578514
step: 200, loss: 0.010706029832363129
step: 210, loss: 0.09043631702661514
step: 220, loss: 0.0031599728390574455
step: 230, loss: 0.001405507093295455
/usr/local/lib/python3.7/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=0.970917225950783, f1=0.9660633484162895,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016912578139454126
step: 10, loss: 0.0011803406523540616
step: 20, loss: 0.0018029235070571303
step: 30, loss: 0.0005872868932783604
step: 40, loss: 0.00047847320092841983
step: 50, loss: 0.00048649648670107126
step: 60, loss: 0.0045272973366081715
step: 70, loss: 0.039292849600315094
step: 80, loss: 0.0006791128544136882
step: 90, loss: 0.02546021342277527

```

```

step: 100, loss: 0.0009001408470794559
step: 110, loss: 0.000566914037335664
step: 120, loss: 0.013899758458137512
step: 130, loss: 0.04119162634015083
step: 140, loss: 0.0026467167772352695
step: 150, loss: 0.013773653656244278
step: 160, loss: 0.0007483020890504122
step: 170, loss: 0.0006179435877129436
step: 180, loss: 0.0008190842345356941
step: 190, loss: 0.00017782581562642008
step: 200, loss: 0.0002238839224446565
step: 210, loss: 0.0017302806954830885
step: 220, loss: 0.00037898815935477614
step: 230, loss: 0.024646075442433357
/usr/local/lib/python3.7/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=0.9699666295884317, f1=0.9617977528089887,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014734531287103891
step: 10, loss: 0.00040915486169978976
step: 20, loss: 0.0003488968650344759
step: 30, loss: 6.672581366728991e-05
step: 40, loss: 0.00027582672191783786
step: 50, loss: 9.075309935724363e-05
step: 60, loss: 0.0003701992682181299
step: 70, loss: 0.13109934329986572
step: 80, loss: 0.0005995313404127955
step: 90, loss: 0.0007480105268768966
step: 100, loss: 0.00024312426103278995
step: 110, loss: 0.0012987449299544096
step: 120, loss: 0.00025772658409550786
step: 130, loss: 0.00034771504579111934
step: 140, loss: 0.00025601740344427526
step: 150, loss: 0.00038021590444259346
step: 160, loss: 8.269817044492811e-05
step: 170, loss: 0.0002215572603745386
step: 180, loss: 0.000651228241622448
step: 190, loss: 0.0003064138873014599
step: 200, loss: 0.009893013164401054
step: 210, loss: 0.0015291848685592413
step: 220, loss: 0.009224781766533852
step: 230, loss: 0.0005820145015604794
/usr/local/lib/python3.7/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=0.9684684684684683, f1=0.9624573378839592,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002163414901588112
step: 10, loss: 0.0008222864125855267
step: 20, loss: 0.0009573324350640178
step: 30, loss: 0.0038673721719533205
step: 40, loss: 0.00023343865177594125
step: 50, loss: 0.0017800785135477781
step: 60, loss: 0.0007627083687111735
step: 70, loss: 0.00017967299208976328
step: 80, loss: 0.00030090659856796265
step: 90, loss: 0.001505669322796166
step: 100, loss: 0.00039980473229661584
step: 110, loss: 0.0002628648071549833
step: 120, loss: 0.00018820667173713446
step: 130, loss: 0.0002114811650244519
step: 140, loss: 0.0007248303736560047
step: 150, loss: 0.0007539950311183929
step: 160, loss: 0.0023013073951005936
step: 170, loss: 0.003825783496722579
step: 180, loss: 0.00043048730003647506
step: 190, loss: 0.0003157739993184805
step: 200, loss: 0.012159788981080055
step: 210, loss: 0.0006662713130936027
step: 220, loss: 0.0005667038494721055
step: 230, loss: 0.0005896365037187934
/usr/local/lib/python3.7/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=0.9706546275395034, f1=0.9648127128263336,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013434139546006918
step: 10, loss: 0.00015825954324100167
step: 20, loss: 0.012956569902598858
step: 30, loss: 0.035452038049697876
step: 40, loss: 0.0002479234535712749
step: 50, loss: 0.0005076837260276079
step: 60, loss: 0.0006490650121122599
step: 70, loss: 0.00025765737518668175
step: 80, loss: 7.043127698125318e-05
step: 90, loss: 0.0019212444312870502

```

```

step: 100, loss: 0.00014524225844070315
step: 110, loss: 0.00014158867998048663
step: 120, loss: 0.0004676530370488763
step: 130, loss: 0.0002232149854535237
step: 140, loss: 0.00013039534678682685
step: 150, loss: 0.00017621327424421906
step: 160, loss: 0.002428739331662655
step: 170, loss: 0.0003938535519409925
step: 180, loss: 0.00013534730533137918
step: 190, loss: 0.000759559334255755
step: 200, loss: 0.00013148627476766706
step: 210, loss: 0.019182734191417694
step: 220, loss: 0.004185180179774761
step: 230, loss: 0.00024027425388339907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9722530521642618, f1=0.9685393258426966,
best_f1=0.9685393258426966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007023802027106285
step: 10, loss: 0.0003214412136003375
step: 20, loss: 0.0002071555791189894
step: 30, loss: 8.81267842487432e-05
step: 40, loss: 0.0009488639771007001
step: 50, loss: 0.0004787073703482747
step: 60, loss: 0.011413905769586563
step: 70, loss: 0.0003912992833647877
step: 80, loss: 0.0006193368462845683
step: 90, loss: 0.00011632448149612173
step: 100, loss: 0.00021218635083641857
step: 110, loss: 0.0003218147321604192
step: 120, loss: 0.0005514424410648644
step: 130, loss: 0.0003732886107172817
step: 140, loss: 0.00019019102910533547
step: 150, loss: 0.00012065715418430045
step: 160, loss: 0.0005071887862868607
step: 170, loss: 0.00025875086430460215
step: 180, loss: 0.013253510929644108
step: 190, loss: 0.0004630870826076716
step: 200, loss: 6.26998589723371e-05
step: 210, loss: 0.0005568534252233803

```

```

step: 220, loss: 0.0001721574371913448
step: 230, loss: 0.001331279636360705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9755555555555556, f1=0.9695603156708005,
best_f1=0.9695603156708005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002237384906038642
step: 10, loss: 0.00035378351458348334
step: 20, loss: 0.0002860235108528286
step: 30, loss: 0.0002714770380407572
step: 40, loss: 0.00034704175777733326
step: 50, loss: 0.00016389155643992126
step: 60, loss: 0.0007029187399893999
step: 70, loss: 0.010952356271445751
step: 80, loss: 0.00014402501983568072
step: 90, loss: 0.0003001003060489893
step: 100, loss: 0.0002658918092492968
step: 110, loss: 0.00021408559405244887
step: 120, loss: 5.228955706115812e-05
step: 130, loss: 0.0006867619813419878
step: 140, loss: 0.00071401905734092
step: 150, loss: 0.0001344732299912721
step: 160, loss: 0.0001716512197162956
step: 170, loss: 0.0003145670925732702
step: 180, loss: 0.00016715087986085564
step: 190, loss: 0.0001234707742696628
step: 200, loss: 0.00022228705347515643
step: 210, loss: 0.0001893881562864408
step: 220, loss: 0.00024127893266268075
step: 230, loss: 0.0006108008092269301
/usr/local/lib/python3.7/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=0.9723145071982282, f1=0.9685393258426966,
best_f1=0.9695603156708005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040416783303953707
step: 10, loss: 0.00014411295705940574

```



```

step: 20, loss: 0.00026751705445349216
step: 30, loss: 0.00019194539345335215
step: 40, loss: 7.579277007607743e-05
step: 50, loss: 0.00012389225594233721
step: 60, loss: 0.0316024012863636
step: 70, loss: 0.0007424521609209478
step: 80, loss: 0.00014607585035264492
step: 90, loss: 8.915123908082023e-05
step: 100, loss: 9.251671872334555e-05
step: 110, loss: 0.00014249687956180423
step: 120, loss: 0.023064542561769485
step: 130, loss: 0.00021236813336145133
step: 140, loss: 0.021852249279618263
step: 150, loss: 0.00041370157850906253
step: 160, loss: 0.0002529648772906512
step: 170, loss: 6.466754712164402e-05
step: 180, loss: 0.00028744328301399946
step: 190, loss: 0.001066901721060276
step: 200, loss: 0.00014660188753623515
step: 210, loss: 0.0017537575913593173
step: 220, loss: 0.00014472482143901289
step: 230, loss: 0.00022146296396385878
/usr/local/lib/python3.7/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=0.9733924611973392, f1=0.9685393258426966,
best_f1=0.9695603156708005
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, 206.40it/s]
load_f1 = 0.9744160177975528
real_f1 = 0.9743016759776536
0it [00:00, ?it/s]/usr/local/lib/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, 188.84it/s]
```

3.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.6121192574501038
```

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

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

```
step: 10, loss: 0.42383185029029846
```

```
step: 20, loss: 0.4018475413322449
```

```
step: 30, loss: 0.35390496253967285
```

```
step: 40, loss: 0.3081027567386627
```

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

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

```
step: 50, loss: 1.081482172012329
```

```
step: 60, loss: 0.19748565554618835
```

```
step: 70, loss: 0.26307690143585205
```

```
step: 80, loss: 0.2541966438293457
```

```
step: 90, loss: 0.368874728679657
```

```
step: 100, loss: 0.25249797105789185
```

```
step: 110, loss: 0.29626473784446716
```

```

step: 120, loss: 0.22598259150981903
step: 130, loss: 0.19215956330299377
step: 140, loss: 0.2878522574901581
step: 150, loss: 0.1268358826637268
step: 160, loss: 0.21349069476127625
step: 170, loss: 0.19139377772808075
step: 180, loss: 0.07286443561315536
step: 190, loss: 0.07002177089452744
step: 200, loss: 0.09549590200185776
step: 210, loss: 0.04082303121685982
step: 220, loss: 0.1125827506184578
step: 230, loss: 0.10328538715839386
step: 240, loss: 0.08727119863033295
step: 250, loss: 0.017483485862612724
step: 260, loss: 0.4036843776702881
step: 270, loss: 0.41500282287597656
step: 280, loss: 0.08120817691087723
step: 290, loss: 0.07833285629749298
step: 300, loss: 0.2892342805862427
step: 310, loss: 0.16989766061306
step: 320, loss: 0.14989088475704193
step: 330, loss: 0.04712636023759842
step: 340, loss: 0.4040769934654236
step: 350, loss: 0.11512477695941925
step: 360, loss: 0.0199830774217844
step: 370, loss: 0.019244063645601273
step: 380, loss: 0.19204865396022797
step: 390, loss: 0.03195849433541298
step: 400, loss: 0.06979138404130936
step: 410, loss: 0.21877512335777283
step: 420, loss: 0.020534304901957512
step: 430, loss: 0.05638428404927254
step: 440, loss: 0.05297653377056122
step: 450, loss: 0.11978878825902939
step: 460, loss: 0.06879697740077972
step: 470, loss: 0.10379204899072647
step: 480, loss: 0.15872913599014282
step: 490, loss: 0.21254697442054749
step: 500, loss: 0.05986650288105011
step: 510, loss: 0.13004277646541595
step: 520, loss: 0.3131929934024811
step: 530, loss: 0.05329997092485428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9214953271028037, f1=0.9166278528178855,
best_f1=0.9166278528178855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08496188372373581
step: 10, loss: 0.05717410519719124
step: 20, loss: 0.09452914446592331
step: 30, loss: 0.1040477454662323
step: 40, loss: 0.09184200316667557
step: 50, loss: 0.055832549929618835
step: 60, loss: 0.06324982643127441
step: 70, loss: 0.06079869344830513
step: 80, loss: 0.034230347722768784
step: 90, loss: 0.050873447209596634
step: 100, loss: 0.1584055870771408
step: 110, loss: 0.02889757603406906
step: 120, loss: 0.08438114821910858
step: 130, loss: 0.00934375450015068
step: 140, loss: 0.09676402807235718
step: 150, loss: 0.03157311677932739
step: 160, loss: 0.045337315648794174
step: 170, loss: 0.04462505504488945
step: 180, loss: 0.09561328589916229
step: 190, loss: 0.013735135085880756
step: 200, loss: 0.29422006011009216
step: 210, loss: 0.03694920614361763
step: 220, loss: 0.012023168615996838
step: 230, loss: 0.09793072193861008
step: 240, loss: 0.025948980823159218
step: 250, loss: 0.04554801434278488
step: 260, loss: 0.05040962994098663
step: 270, loss: 0.09989520162343979
step: 280, loss: 0.020698966458439827
step: 290, loss: 0.0289317574352026
step: 300, loss: 0.09494030475616455
step: 310, loss: 0.10307823121547699
step: 320, loss: 0.09819109737873077
step: 330, loss: 0.08138688653707504
step: 340, loss: 0.13082970678806305
step: 350, loss: 0.007083873264491558
step: 360, loss: 0.09053433686494827
step: 370, loss: 0.01391814835369587
step: 380, loss: 0.1305844783782959
step: 390, loss: 0.012906328774988651
step: 400, loss: 0.14097581803798676
step: 410, loss: 0.020651767030358315
```

```

step: 420, loss: 0.015089581720530987
step: 430, loss: 0.08855319768190384
step: 440, loss: 0.0029500266537070274
step: 450, loss: 0.052246108651161194
step: 460, loss: 0.0947151631116867
step: 470, loss: 0.040131840854883194
step: 480, loss: 0.03986596316099167
step: 490, loss: 0.05084920674562454
step: 500, loss: 0.004417087882757187
step: 510, loss: 0.056120142340660095
step: 520, loss: 0.22556063532829285
step: 530, loss: 0.023978402838110924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9373865698729582, f1=0.9300635785649409,
best_f1=0.9300635785649409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07363659888505936
step: 10, loss: 0.10496219247579575
step: 20, loss: 0.010887200944125652
step: 30, loss: 0.09259909391403198
step: 40, loss: 0.02301725000143051
step: 50, loss: 0.022877776995301247
step: 60, loss: 0.07807255536317825
step: 70, loss: 0.005092364735901356
step: 80, loss: 0.044785283505916595
step: 90, loss: 0.007367152255028486
step: 100, loss: 0.08049412816762924
step: 110, loss: 0.008773528970777988
step: 120, loss: 0.1155741810798645
step: 130, loss: 0.06908275932073593
step: 140, loss: 0.02411993220448494
step: 150, loss: 0.018329543992877007
step: 160, loss: 0.05302412062883377
step: 170, loss: 0.014052332378923893
step: 180, loss: 0.015233148820698261
step: 190, loss: 0.004904359579086304
step: 200, loss: 0.07800590246915817
step: 210, loss: 0.03480119630694389
step: 220, loss: 0.12434211373329163
step: 230, loss: 0.10395289957523346

```

```

step: 240, loss: 0.027949465438723564
step: 250, loss: 0.17184486985206604
step: 260, loss: 0.12331783771514893
step: 270, loss: 0.008395815268158913
step: 280, loss: 0.09796316921710968
step: 290, loss: 0.016069939360022545
step: 300, loss: 0.08567101508378983
step: 310, loss: 0.11418188363313675
step: 320, loss: 0.0223813746124506
step: 330, loss: 0.01130672451108694
step: 340, loss: 0.02927933633327484
step: 350, loss: 0.05965791270136833
step: 360, loss: 0.009800823405385017
step: 370, loss: 0.10381091386079788
step: 380, loss: 0.05055089667439461
step: 390, loss: 0.030086655169725418
step: 400, loss: 0.04036877304315567
step: 410, loss: 0.06768734008073807
step: 420, loss: 0.007823643274605274
step: 430, loss: 0.014793409034609795
step: 440, loss: 0.17779584228992462
step: 450, loss: 0.040730010718107224
step: 460, loss: 0.0337701141834259
step: 470, loss: 0.07651641219854355
step: 480, loss: 0.16448649764060974
step: 490, loss: 0.008429172448813915
step: 500, loss: 0.010566038079559803
step: 510, loss: 0.10444482415914536
step: 520, loss: 0.006797553040087223
step: 530, loss: 0.00791106652468443
/usr/local/lib/python3.7/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=0.9277389277389277, f1=0.9222170470423847,
best_f1=0.9300635785649409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01343084778636694
step: 10, loss: 0.0009579880279488862
step: 20, loss: 0.08054955303668976
step: 30, loss: 0.12627862393856049
step: 40, loss: 0.00622148672118783
step: 50, loss: 0.03493181988596916
step: 60, loss: 0.013549137860536575
step: 70, loss: 0.08344827592372894
step: 80, loss: 0.06030815467238426
step: 90, loss: 0.007373876869678497

```

```
step: 100, loss: 0.0015952590620145202
step: 110, loss: 0.15165942907333374
step: 120, loss: 0.0023252065293490887
step: 130, loss: 0.022572610527276993
step: 140, loss: 0.02731209248304367
step: 150, loss: 0.003990980796515942
step: 160, loss: 0.004324426408857107
step: 170, loss: 0.04755097255110741
step: 180, loss: 0.04322529956698418
step: 190, loss: 0.1252419352531433
step: 200, loss: 0.1191127821803093
step: 210, loss: 0.006191943772137165
step: 220, loss: 0.032494280487298965
step: 230, loss: 0.01744864135980606
step: 240, loss: 0.01718064583837986
step: 250, loss: 0.07117177546024323
step: 260, loss: 0.003154327627271414
step: 270, loss: 0.07331694662570953
step: 280, loss: 0.0022404545452445745
step: 290, loss: 0.051078975200653076
step: 300, loss: 0.0016763542080298066
step: 310, loss: 0.00981905497610569
step: 320, loss: 0.046965863555669785
step: 330, loss: 0.03563208132982254
step: 340, loss: 0.0037323576398193836
step: 350, loss: 0.02539677359163761
step: 360, loss: 0.04251078888773918
step: 370, loss: 0.0011775284074246883
step: 380, loss: 0.0037279007956385612
step: 390, loss: 0.0010679494589567184
step: 400, loss: 0.004594224039465189
step: 410, loss: 0.0078062862157821655
step: 420, loss: 0.037429142743349075
step: 430, loss: 0.03599116951227188
step: 440, loss: 0.012781974859535694
step: 450, loss: 0.05346385017037392
step: 460, loss: 0.010463116690516472
step: 470, loss: 0.0067536234855651855
step: 480, loss: 0.039160050451755524
step: 490, loss: 0.01014744769781828
step: 500, loss: 0.08450981229543686
step: 510, loss: 0.09889407455921173
step: 520, loss: 0.002450640080496669
step: 530, loss: 0.1388627290725708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```


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=0.9400369003690037, f1=0.9310187300137049,

best_f1=0.9310187300137049

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

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

step: 0, loss: 0.002522457856684923

step: 10, loss: 0.01616646721959114

step: 20, loss: 0.08894730359315872

step: 30, loss: 0.006280353758484125

step: 40, loss: 0.0005105827003717422

step: 50, loss: 0.022428618744015694

step: 60, loss: 0.01558894943445921

step: 70, loss: 0.002181706018745899

step: 80, loss: 0.016587018966674805

step: 90, loss: 0.1984703093767166

step: 100, loss: 0.004209378734230995

step: 110, loss: 0.011916394345462322

step: 120, loss: 0.0305988397449255

step: 130, loss: 0.019204525277018547

step: 140, loss: 0.007806312758475542

step: 150, loss: 0.12035978585481644

step: 160, loss: 0.037487003952264786

step: 170, loss: 0.01964130438864231

step: 180, loss: 0.007431862410157919

step: 190, loss: 0.0011356838513165712

step: 200, loss: 0.004215225111693144

step: 210, loss: 0.0005827798740938306

step: 220, loss: 0.0031643377151340246

step: 230, loss: 0.002510777907446027

step: 240, loss: 0.0020220146980136633

step: 250, loss: 0.15500353276729584

step: 260, loss: 0.0006104707135818899

step: 270, loss: 0.0002799765788950026

step: 280, loss: 0.0044565764255821705

step: 290, loss: 0.03224070742726326

step: 300, loss: 0.014321713708341122

step: 310, loss: 0.010079235769808292

step: 320, loss: 0.07352221757173538

step: 330, loss: 0.0038160434924066067

step: 340, loss: 0.00853133387863636

step: 350, loss: 0.0019875362049788237

step: 360, loss: 0.0319281704723835

step: 370, loss: 0.002450184430927038

step: 380, loss: 0.0004544180410448462

step: 390, loss: 0.014887957833707333

```

step: 400, loss: 0.0004946582484990358
step: 410, loss: 0.17127101123332977
step: 420, loss: 0.13646067678928375
step: 430, loss: 0.013229488395154476
step: 440, loss: 0.0017397733172401786
step: 450, loss: 0.023919932544231415
step: 460, loss: 0.013671970926225185
step: 470, loss: 0.08558087050914764
step: 480, loss: 0.002708561485633254
step: 490, loss: 0.02680949494242668
step: 500, loss: 0.011494322679936886
step: 510, loss: 0.010127742774784565
step: 520, loss: 0.11589761823415756
step: 530, loss: 0.02244153805077076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9407372841810545, f1=0.9343880874825501,
best_f1=0.9343880874825501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011010624468326569
step: 10, loss: 0.00038944536936469376
step: 20, loss: 0.0005676791188307106
step: 30, loss: 0.0012978068552911282
step: 40, loss: 0.00043414102401584387
step: 50, loss: 0.000255197926890105
step: 60, loss: 0.0005400898517109454
step: 70, loss: 0.0008055094513110816
step: 80, loss: 0.0016621637623757124
step: 90, loss: 0.0005409475415945053
step: 100, loss: 0.028227275237441063
step: 110, loss: 0.008677211590111256
step: 120, loss: 0.00970855075865984
step: 130, loss: 0.002293655415996909
step: 140, loss: 0.008790642954409122
step: 150, loss: 0.028190908953547478
step: 160, loss: 0.04836675152182579
step: 170, loss: 0.014327767305076122
step: 180, loss: 0.0043853032402694225
step: 190, loss: 0.06047752499580383
step: 200, loss: 0.010759216733276844
step: 210, loss: 0.0036809127777814865

```

```

step: 220, loss: 0.003276368835940957
step: 230, loss: 0.0913114994764328
step: 240, loss: 0.01007112767547369
step: 250, loss: 0.07905932515859604
step: 260, loss: 0.002774938242509961
step: 270, loss: 0.002083136234432459
step: 280, loss: 0.03460419923067093
step: 290, loss: 0.10803956538438797
step: 300, loss: 0.0223095640540123
step: 310, loss: 0.07008273899555206
step: 320, loss: 0.0005976933753117919
step: 330, loss: 0.0004508448764681816
step: 340, loss: 0.0026238905265927315
step: 350, loss: 0.0008070504991337657
step: 360, loss: 0.007266020867973566
step: 370, loss: 0.001895554712973535
step: 380, loss: 0.0011692434782162309
step: 390, loss: 0.0025176263879984617
step: 400, loss: 0.12056942284107208
step: 410, loss: 0.017123673111200333
step: 420, loss: 0.00235170079395175
step: 430, loss: 0.0010778596624732018
step: 440, loss: 0.0008815351175144315
step: 450, loss: 0.2754954695701599
step: 460, loss: 0.00483269989490509
step: 470, loss: 0.008815971203148365
step: 480, loss: 0.012411898002028465
step: 490, loss: 0.07722488790750504
step: 500, loss: 0.056721415370702744
step: 510, loss: 0.10088478773832321
step: 520, loss: 0.004257837310433388
step: 530, loss: 0.0026875585317611694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9410125406409661, f1=0.9288354898336414,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008306547068059444
step: 10, loss: 0.007315136957913637
step: 20, loss: 0.0016214123461395502
step: 30, loss: 0.00924109946936369

```

step: 40, loss: 0.0038155755028128624
step: 50, loss: 0.021822800859808922
step: 60, loss: 0.011458100751042366
step: 70, loss: 0.006381336133927107
step: 80, loss: 0.02496090531349182
step: 90, loss: 0.0006121046608313918
step: 100, loss: 0.0006389784393832088
step: 110, loss: 0.0002281538036186248
step: 120, loss: 0.0046987575478851795
step: 130, loss: 0.00020543417485896498
step: 140, loss: 0.00017579187988303602
step: 150, loss: 0.0009403462754562497
step: 160, loss: 0.0013518871273845434
step: 170, loss: 0.010428954847157001
step: 180, loss: 0.01128404401242733
step: 190, loss: 0.023624615743756294
step: 200, loss: 0.006690707989037037
step: 210, loss: 0.017069173976778984
step: 220, loss: 0.023691242560744286
step: 230, loss: 0.008301756344735622
step: 240, loss: 0.025621086359024048
step: 250, loss: 0.01803617924451828
step: 260, loss: 0.00017951949848793447
step: 270, loss: 0.0014998730039224029
step: 280, loss: 0.00738248135894537
step: 290, loss: 0.0006805713637731969
step: 300, loss: 0.00018582721531856805
step: 310, loss: 0.000590464158449322
step: 320, loss: 0.0007458623149432242
step: 330, loss: 6.711120659019798e-05
step: 340, loss: 5.873427653568797e-05
step: 350, loss: 0.04812256246805191
step: 360, loss: 0.0013343766331672668
step: 370, loss: 0.0004286908369977027
step: 380, loss: 0.0101162726059556
step: 390, loss: 0.0009006822947412729
step: 400, loss: 0.01237599365413189
step: 410, loss: 0.00034499066532589495
step: 420, loss: 0.07226074486970901
step: 430, loss: 0.0009844573214650154
step: 440, loss: 0.0022226315923035145
step: 450, loss: 0.002455400303006172
step: 460, loss: 0.0007913429290056229
step: 470, loss: 0.21587561070919037
step: 480, loss: 0.006030368618667126
step: 490, loss: 0.006206311751157045
step: 500, loss: 0.002383613958954811
step: 510, loss: 0.0028534233570098877

```

step: 520, loss: 0.0006144416402094066
step: 530, loss: 0.0005563702434301376
/usr/local/lib/python3.7/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=0.935813953488372, f1=0.9298653042266605,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005519850528798997
step: 10, loss: 0.0008871410973370075
step: 20, loss: 0.0008712325361557305
step: 30, loss: 0.024590803310275078
step: 40, loss: 0.00019741302821785212
step: 50, loss: 0.005123603157699108
step: 60, loss: 0.001322206575423479
step: 70, loss: 0.0009121911134570837
step: 80, loss: 0.0002227222139481455
step: 90, loss: 0.00029389746487140656
step: 100, loss: 0.0005061595584265888
step: 110, loss: 0.0007264123414643109
step: 120, loss: 0.0005701036425307393
step: 130, loss: 0.0480586439371109
step: 140, loss: 0.000961028563324362
step: 150, loss: 0.0017629426438361406
step: 160, loss: 0.0003014874819200486
step: 170, loss: 0.07394660264253616
step: 180, loss: 0.003339895512908697
step: 190, loss: 0.06389180570840836
step: 200, loss: 0.0013831857359036803
step: 210, loss: 0.03376486524939537
step: 220, loss: 0.03176935762166977
step: 230, loss: 0.016614526510238647
step: 240, loss: 0.006578702945262194
step: 250, loss: 4.16661350755021e-05
step: 260, loss: 0.00038849611883051693
step: 270, loss: 0.011968731880187988
step: 280, loss: 0.0018908651545643806
step: 290, loss: 0.025431472808122635
step: 300, loss: 0.0004365986096672714
step: 310, loss: 0.0034321737475693226
step: 320, loss: 0.0004752209933940321
step: 330, loss: 0.00014478246157523245
step: 340, loss: 0.009850356727838516
step: 350, loss: 0.018375037238001823
step: 360, loss: 0.0005555941606871784
step: 370, loss: 0.11012028157711029

```

```

step: 380, loss: 0.0004643301072064787
step: 390, loss: 0.0010630181059241295
step: 400, loss: 0.011295522563159466
step: 410, loss: 0.0065159061923623085
step: 420, loss: 0.0014077594969421625
step: 430, loss: 0.027615271508693695
step: 440, loss: 0.0037882912438362837
step: 450, loss: 0.001045743003487587
step: 460, loss: 0.009683040902018547
step: 470, loss: 0.055761393159627914
step: 480, loss: 0.00033487126347608864
step: 490, loss: 0.0058087012730538845
step: 500, loss: 0.0001454894372727722
step: 510, loss: 0.005773133132606745
step: 520, loss: 0.0009906365303322673
step: 530, loss: 0.0009541079634800553
/usr/local/lib/python3.7/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=0.9266943291839558, f1=0.9203703703703704,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014425076369661838
step: 10, loss: 0.00224858894944191
step: 20, loss: 0.0006576580926775932
step: 30, loss: 0.005415312480181456
step: 40, loss: 0.0006986422231420875
step: 50, loss: 0.0007504284149035811
step: 60, loss: 0.0007089728605933487
step: 70, loss: 0.005060738883912563
step: 80, loss: 0.0012964048655703664
step: 90, loss: 0.0940164178609848
step: 100, loss: 0.0007476522587239742
step: 110, loss: 0.004115724936127663
step: 120, loss: 0.0023275220301002264
step: 130, loss: 0.001633281703107059
step: 140, loss: 0.0015344441635534167
step: 150, loss: 0.0022151079028844833
step: 160, loss: 0.0007072381558828056
step: 170, loss: 0.0006615976453758776
step: 180, loss: 0.0005226685316301882
step: 190, loss: 0.0003296601353213191
step: 200, loss: 0.0003441354201640934
step: 210, loss: 0.0009632020955905318
step: 220, loss: 0.0002923124411609024
step: 230, loss: 0.0009645267273299396

```

```

step: 240, loss: 0.002178015885874629
step: 250, loss: 0.00031641797977499664
step: 260, loss: 8.05236486485228e-05
step: 270, loss: 0.00015068324864841998
step: 280, loss: 0.00017978403775487095
step: 290, loss: 0.0001302196178585291
step: 300, loss: 0.00413889903575182
step: 310, loss: 0.05788553133606911
step: 320, loss: 9.17876313906163e-05
step: 330, loss: 0.00018499490397516638
step: 340, loss: 0.0026912030298262835
step: 350, loss: 0.0324963741004467
step: 360, loss: 0.0003335595247335732
step: 370, loss: 0.00025089780683629215
step: 380, loss: 0.005844457074999809
step: 390, loss: 8.628334035165608e-05
step: 400, loss: 0.02478015050292015
step: 410, loss: 0.0012882197042927146
step: 420, loss: 0.0011919753160327673
step: 430, loss: 0.00019489394617266953
step: 440, loss: 0.00015160816838033497
step: 450, loss: 0.00014387437840923667
step: 460, loss: 0.00010413698328193277
step: 470, loss: 6.967712397454306e-05
step: 480, loss: 0.0010109470458701253
step: 490, loss: 0.0004538570065051317
step: 500, loss: 0.0013327104970812798
step: 510, loss: 0.012067453935742378
step: 520, loss: 0.027243221178650856
step: 530, loss: 0.0049399579875171185
/usr/local/lib/python3.7/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=0.9354389224338133, f1=0.9312413474850022,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006384536158293486
step: 10, loss: 0.0010572567116469145
step: 20, loss: 0.016828766092658043
step: 30, loss: 0.008405669592320919
step: 40, loss: 9.214784222422168e-05
step: 50, loss: 0.00034823184250853956
step: 60, loss: 0.014047103002667427
step: 70, loss: 0.00018617087334860116
step: 80, loss: 0.00010199628013651818
step: 90, loss: 8.629407966509461e-05

```

```
step: 100, loss: 0.0008287926320917904
step: 110, loss: 0.006167355924844742
step: 120, loss: 0.0028053068090230227
step: 130, loss: 0.028285695239901543
step: 140, loss: 0.0018146611982956529
step: 150, loss: 0.001264976803213358
step: 160, loss: 0.07710138708353043
step: 170, loss: 0.0006490636733360589
step: 180, loss: 0.00036698434269055724
step: 190, loss: 0.0006725152488797903
step: 200, loss: 0.012073975056409836
step: 210, loss: 0.00041981006506830454
step: 220, loss: 0.0002158879506168887
step: 230, loss: 0.00045900107943452895
step: 240, loss: 0.0006303437403403223
step: 250, loss: 0.012383250519633293
step: 260, loss: 0.00371731910854578
step: 270, loss: 0.0005082353018224239
step: 280, loss: 0.0021873535588383675
step: 290, loss: 0.00043086739606224
step: 300, loss: 0.0014041932299733162
step: 310, loss: 0.003940865863114595
step: 320, loss: 0.0003908926446456462
step: 330, loss: 0.0017712571425363421
step: 340, loss: 0.000270684773568064
step: 350, loss: 0.00019762157171498984
step: 360, loss: 0.00997441541403532
step: 370, loss: 0.0042059277184307575
step: 380, loss: 0.0005060954135842621
step: 390, loss: 0.00012978199811186641
step: 400, loss: 0.00015625533706042916
step: 410, loss: 0.00019252942001912743
step: 420, loss: 0.00012424946180544794
step: 430, loss: 7.418156019411981e-05
step: 440, loss: 0.00010525410471018404
step: 450, loss: 0.010684613138437271
step: 460, loss: 8.41641885926947e-05
step: 470, loss: 0.0015148509992286563
step: 480, loss: 0.00012234217138029635
step: 490, loss: 0.0012450431240722537
step: 500, loss: 0.00019290427735541016
step: 510, loss: 0.00020513232448138297
step: 520, loss: 0.0003041772579308599
step: 530, loss: 0.0032064293045550585
/usr/local/lib/python3.7/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=0.9384469696969696, f1=0.9303647560397915,
```



```
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016090032295323908
step: 10, loss: 0.00013546633999794722
step: 20, loss: 0.000267073541181162
step: 30, loss: 0.00017187136108987033
step: 40, loss: 9.547971421852708e-05
step: 50, loss: 9.676415356807411e-05
step: 60, loss: 0.00011746268137358129
step: 70, loss: 0.00024978964938782156
step: 80, loss: 0.00015994074055925012
step: 90, loss: 0.00031967050745151937
step: 100, loss: 0.0015730186132714152
step: 110, loss: 0.00019511335995048285
step: 120, loss: 0.015197858214378357
step: 130, loss: 9.116787259699777e-05
step: 140, loss: 0.00010992953320965171
step: 150, loss: 0.009368409402668476
step: 160, loss: 0.0009449404897168279
step: 170, loss: 0.0002872862678486854
step: 180, loss: 0.00031683719134889543
step: 190, loss: 0.0004386872751638293
step: 200, loss: 0.014847640879452229
step: 210, loss: 0.009078709408640862
step: 220, loss: 0.00489412946626544
step: 230, loss: 4.394038478494622e-05
step: 240, loss: 0.0005268228123895824
step: 250, loss: 0.0004582985129673034
step: 260, loss: 0.0033265724778175354
step: 270, loss: 0.0003000361903104931
step: 280, loss: 0.001114025479182601
step: 290, loss: 0.0002701566554605961
step: 300, loss: 0.00019219038949813694
step: 310, loss: 0.014130097813904285
step: 320, loss: 0.0001792058174032718
step: 330, loss: 0.00017288528033532202
step: 340, loss: 0.0003998038300778717
step: 350, loss: 0.00013326997577678412
step: 360, loss: 0.00023814887390471995
step: 370, loss: 0.0025778799317777157
step: 380, loss: 0.0009386194869875908
step: 390, loss: 0.0005906407604925334
step: 400, loss: 0.00033017914392985404
step: 410, loss: 0.0007869397522881627
step: 420, loss: 5.668135418090969e-05
step: 430, loss: 0.0006807839381508529
```

```

step: 440, loss: 0.00014497477968689054
step: 450, loss: 0.0007542610401287675
step: 460, loss: 0.0002418391959508881
step: 470, loss: 0.00015774993516970426
step: 480, loss: 8.648385846754536e-05
step: 490, loss: 0.01380988396704197
step: 500, loss: 8.68890929268673e-05
step: 510, loss: 8.474772766930982e-05
step: 520, loss: 0.0001047837213263847
step: 530, loss: 8.11096397228539e-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=0.9345099860659545, f1=0.9284386617100371,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014864432159811258
step: 10, loss: 0.0011785810347646475
step: 20, loss: 0.0004840435285586864
step: 30, loss: 4.312378223403357e-05
step: 40, loss: 0.00011452742910478264
step: 50, loss: 7.794434350216761e-05
step: 60, loss: 0.0002885498106479645
step: 70, loss: 0.0008053105557337403
step: 80, loss: 0.00010820886382134631
step: 90, loss: 0.0009525057394057512
step: 100, loss: 0.23755693435668945
step: 110, loss: 0.00037675967905670404
step: 120, loss: 0.002888987772166729
step: 130, loss: 0.0015660749049857259
step: 140, loss: 0.0003121968766208738
step: 150, loss: 0.0028313305228948593
step: 160, loss: 0.0003428765630815178
step: 170, loss: 0.00034360826248303056
step: 180, loss: 0.0009473403333686292
step: 190, loss: 0.0003541400656104088
step: 200, loss: 0.001393306301906705
step: 210, loss: 0.00017404588288627565
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.013300190679728985
step: 230, loss: 0.0008503867429681122
step: 240, loss: 0.0008275664877146482
step: 250, loss: 0.00010503988596610725
step: 260, loss: 0.0001209435286000371
step: 270, loss: 0.0030963323079049587
step: 280, loss: 0.0006381715065799654

```

```

step: 290, loss: 0.00365203688852489
step: 300, loss: 0.0005970275960862637
step: 310, loss: 0.00025059570907615125
step: 320, loss: 0.0002636900171637535
step: 330, loss: 0.0005113529623486102
step: 340, loss: 0.019720738753676414
step: 350, loss: 0.0002119874843629077
step: 360, loss: 0.0007377708097919822
step: 370, loss: 0.0004944014362990856
step: 380, loss: 0.0001070220532710664
step: 390, loss: 0.00030206769588403404
step: 400, loss: 0.0015207031974568963
step: 410, loss: 0.012338199652731419
step: 420, loss: 0.00032411314896307886
step: 430, loss: 0.0008926903828978539
step: 440, loss: 0.0004379782185424119
step: 450, loss: 0.0014993313234299421
step: 460, loss: 0.00013681068958248943
step: 470, loss: 0.0019080708734691143
step: 480, loss: 0.00014921798720024526
step: 490, loss: 0.0006377403624355793
step: 500, loss: 0.00015115532733034343
step: 510, loss: 0.00014742308121640235
step: 520, loss: 0.0022426315117627382
step: 530, loss: 0.008160787634551525
/usr/local/lib/python3.7/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=0.9399720800372265, f1=0.934640522875817,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001040478891809471
step: 10, loss: 0.0012886926997452974
step: 20, loss: 0.00023550693003926426
step: 30, loss: 7.787660433677956e-05
step: 40, loss: 8.492953929817304e-05
step: 50, loss: 0.0005304103251546621
step: 60, loss: 0.0002766066463664174
step: 70, loss: 0.024047154933214188
step: 80, loss: 0.00019379211880732328
step: 90, loss: 0.00010340099106542766
step: 100, loss: 9.192756260745227e-05
step: 110, loss: 5.804605461889878e-05
step: 120, loss: 0.0001091526064556092
step: 130, loss: 4.466372774913907e-05
step: 140, loss: 0.00030795729253441095

```

```

step: 150, loss: 0.0003210896102245897
step: 160, loss: 0.00018113885016646236
step: 170, loss: 0.00030473171500489116
step: 180, loss: 0.00010174237104365602
step: 190, loss: 0.00010497013136046007
step: 200, loss: 8.09489720268175e-05
step: 210, loss: 7.460310735041276e-05
step: 220, loss: 8.495996735291556e-05
step: 230, loss: 0.0003287687723059207
step: 240, loss: 0.0005944420117884874
step: 250, loss: 0.0002691995177883655
step: 260, loss: 9.436573600396514e-05
step: 270, loss: 0.06791772693395615
step: 280, loss: 0.008546056225895882
step: 290, loss: 5.174109901417978e-05
step: 300, loss: 7.463446672772989e-05
step: 310, loss: 8.32212608656846e-05
step: 320, loss: 0.00012154960131738335
step: 330, loss: 0.0001046962570399046
step: 340, loss: 0.004867462441325188
step: 350, loss: 5.216518184170127e-05
step: 360, loss: 0.12114015966653824
step: 370, loss: 6.696938362438232e-05
step: 380, loss: 0.00013451107952278107
step: 390, loss: 0.00011411661398597062
step: 400, loss: 0.00019598493236117065
step: 410, loss: 0.0006531640538014472
step: 420, loss: 8.053966303123161e-05
step: 430, loss: 0.00012753727787639946
step: 440, loss: 6.178251351229846e-05
step: 450, loss: 0.00019358993449714035
step: 460, loss: 0.1492377519607544
step: 470, loss: 0.0005784914246760309
step: 480, loss: 4.4040159991709515e-05
step: 490, loss: 0.006098778452724218
step: 500, loss: 4.655395241570659e-05
step: 510, loss: 0.0003650221333373338
step: 520, loss: 0.00017998529074247926
step: 530, loss: 0.0001433767902199179
/usr/local/lib/python3.7/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=0.9374999999999999, f1=0.932573599240266,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040902773616835475

```

step: 10, loss: 0.0001652316568652168
step: 20, loss: 0.00021855100931134075
step: 30, loss: 0.00020219414727762341
step: 40, loss: 0.00038611303898505867
step: 50, loss: 0.00031995863537304103
step: 60, loss: 0.00017488410230726004
step: 70, loss: 0.00019198944210074842
step: 80, loss: 0.00021705124527215958
step: 90, loss: 0.0001312886452069506
step: 100, loss: 0.0001308432110818103
step: 110, loss: 0.00023178133415058255
step: 120, loss: 0.0001557897630846128
step: 130, loss: 0.00014984860899858177
step: 140, loss: 0.0001433887955499813
step: 150, loss: 0.0004228127363603562
step: 160, loss: 0.0003075350832659751
step: 170, loss: 0.00011903386621270329
step: 180, loss: 9.666284313425422e-05
step: 190, loss: 0.003026213962584734
step: 200, loss: 0.0001583422563271597
step: 210, loss: 0.0025928346440196037
step: 220, loss: 0.00014462748367805034
step: 230, loss: 0.00019286057795397937
step: 240, loss: 0.0011306521482765675
step: 250, loss: 0.00018381114932708442
step: 260, loss: 0.00023904630506876856
step: 270, loss: 0.0009072681423276663
step: 280, loss: 0.00022390027879737318
step: 290, loss: 0.00016282680735457689
step: 300, loss: 0.01397226843982935
step: 310, loss: 0.0009130730177275836
step: 320, loss: 0.0003504786582197994
step: 330, loss: 0.00022387923672795296
step: 340, loss: 0.0008860588422976434
step: 350, loss: 0.00021618443133775145
step: 360, loss: 0.00018290430307388306
step: 370, loss: 0.0002847051073331386
step: 380, loss: 0.000458572234492749
step: 390, loss: 0.00010398713493486866
step: 400, loss: 0.0018765468848869205
step: 410, loss: 8.155130490195006e-05
step: 420, loss: 0.00013210027827881277
step: 430, loss: 0.00014845943951513618
step: 440, loss: 0.00012750981841236353
step: 450, loss: 0.00012660866195801646
step: 460, loss: 0.034887973219156265
step: 470, loss: 0.00018928565259557217
step: 480, loss: 0.0001368463272228837

```

step: 490, loss: 0.00017752843268681318
step: 500, loss: 0.0013041467173025012
step: 510, loss: 0.02585318125784397
step: 520, loss: 5.0428923714207485e-05
step: 530, loss: 0.00030340204830281436
/usr/local/lib/python3.7/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=0.9380281690140844, f1=0.9359698681732581,
best_f1=0.9288354898336414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011272125993855298
step: 10, loss: 0.0006306098657660186
step: 20, loss: 0.00033057096879929304
step: 30, loss: 0.00011397384514566511
step: 40, loss: 0.00015335889474954456
step: 50, loss: 0.0008270300459116697
step: 60, loss: 0.0002667838998604566
step: 70, loss: 0.0001295155379921198
step: 80, loss: 0.00012919188884552568
step: 90, loss: 8.8360087829642e-05
step: 100, loss: 3.055315028177574e-05
step: 110, loss: 0.0001331671664956957
step: 120, loss: 8.446873107459396e-05
step: 130, loss: 0.00011010146408807486
step: 140, loss: 0.0001182620762847364
step: 150, loss: 0.00018297290080226958
step: 160, loss: 0.0001484598615206778
step: 170, loss: 6.032901364960708e-05
step: 180, loss: 0.00020048744045197964
step: 190, loss: 0.007922794669866562
step: 200, loss: 0.00024197019229177386
step: 210, loss: 5.919438990531489e-05
step: 220, loss: 0.00016617804067209363
step: 230, loss: 0.0001780721067916602
step: 240, loss: 0.00012450867507141083
step: 250, loss: 7.66603261581622e-05
step: 260, loss: 0.00010206326260231435
step: 270, loss: 0.0001638523826841265
step: 280, loss: 0.00029155731317587197
step: 290, loss: 6.31563161732629e-05
step: 300, loss: 0.0001141685206675902
step: 310, loss: 0.010035504586994648
step: 320, loss: 0.00015899150457698852
step: 330, loss: 0.0001693834492471069
step: 340, loss: 0.00023490242892876267

```

```

step: 350, loss: 0.015259936451911926
step: 360, loss: 0.00023656297707930207
step: 370, loss: 0.00023505405988544226
step: 380, loss: 0.00014498207019641995
step: 390, loss: 0.00013398278679233044
step: 400, loss: 0.00020296548609621823
step: 410, loss: 0.0001703360176179558
step: 420, loss: 0.00016577223141212016
step: 430, loss: 0.00010981170635204762
step: 440, loss: 0.00012083345063729212
step: 450, loss: 0.0014530429616570473
step: 460, loss: 0.00018429706688039005
step: 470, loss: 0.00013592375034932047
step: 480, loss: 0.006197161506861448
step: 490, loss: 0.00012167299428256229
step: 500, loss: 0.00025689712492749095
step: 510, loss: 0.00011264430213486776
step: 520, loss: 0.0001589688181411475
step: 530, loss: 7.546960841864347e-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=0.9382600561272217, f1=0.9353327085285849,
best_f1=0.9288354898336414
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 [00:23, 247.36it/s]
load_f1 = 0.9387755102040817
real_f1 = 0.9373549883990719
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 195.32it/s]

```

4 DITTO

4.1 DITTO STRUCTURED

4.1.1 Beer - Running the matcher

```

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

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

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \

```



```

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 372kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 809kB/s]
Downloading: 100% 456k/456k [00:01<00:00, 399kB/s]
Downloading: 100% 501M/501M [00:11<00:00, 42.5MB/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: 0.4315553903579712
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=0.3157894736842105, f1=0.34666666666666666,
best_f1=0.34666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45601508021354675
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.358974358974359, f1=0.35135135135135137,
best_f1=0.35135135135135137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4076765477657318
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.38095238095238093, f1=0.2857142857142857,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2032277137041092
/usr/local/lib/python3.7/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=0.3023255813953489,
best_f1=0.2857142857142857

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3202267289161682
/usr/local/lib/python3.7/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=0.34146341463414637, f1=0.3023255813953489,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2587318420410156
/usr/local/lib/python3.7/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=0.32000000000000006, f1=0.33766233766233766,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6481027603149414
/usr/local/lib/python3.7/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=0.30434782608695654, f1=0.2826086956521739,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5202639698982239
/usr/local/lib/python3.7/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=0.29166666666666667, f1=0.27368421052631575,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4113464057445526
/usr/local/lib/python3.7/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=0.3010752688172043, f1=0.28571428571428575,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.47326239943504333
/usr/local/lib/python3.7/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=0.32911392405063294, f1=0.3170731707317073,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45237696170806885
/usr/local/lib/python3.7/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=0.32098765432098764, f1=0.3170731707317073,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33966776728630066
/usr/local/lib/python3.7/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=0.3225806451612903, f1=0.16666666666666666,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3894004225730896
/usr/local/lib/python3.7/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=0.3783783783783784, f1=0.3888888888888889,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.356864869594574
/usr/local/lib/python3.7/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=0.3783783783783784, f1=0.3888888888888889,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39496102929115295
/usr/local/lib/python3.7/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=0.3783783783783784, f1=0.3888888888888889,
best_f1=0.2857142857142857
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, 113697.25it/s]
/usr/local/lib/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 = 0.31578947368421056
real_f1 = 0.3188405797101449
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.81it/s]

```

4.1.2 DBLP-ACM - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16

```

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.588969886302948
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4717990458011627
step: 20, loss: 0.5044203996658325
step: 30, loss: 0.3129245340824127
step: 40, loss: 0.36869847774505615
step: 50, loss: 0.5156593918800354
step: 60, loss: 0.25395193696022034
step: 70, loss: 0.20164331793785095
step: 80, loss: 0.22155353426933289
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.2882879674434662
step: 100, loss: 0.13341660797595978
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 110, loss: 0.2317131757736206
step: 120, loss: 0.1680179387331009
step: 130, loss: 0.09873668104410172
step: 140, loss: 0.005864291451871395
step: 150, loss: 0.03798158839344978
step: 160, loss: 0.04644940048456192
step: 170, loss: 0.045027486979961395
step: 180, loss: 0.017319386824965477
step: 190, loss: 0.014497119002044201
step: 200, loss: 0.058654699474573135
step: 210, loss: 0.003924332093447447
step: 220, loss: 0.03829212114214897
step: 230, loss: 0.010932579636573792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9710467706013363, f1=0.967525195968645,
best_f1=0.967525195968645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00898749940097332
step: 10, loss: 0.08090133219957352
step: 20, loss: 0.02830246463418007
step: 30, loss: 0.3061143159866333
step: 40, loss: 0.058999866247177124
step: 50, loss: 0.009409312158823013
step: 60, loss: 0.006707961205393076
step: 70, loss: 0.13197416067123413
step: 80, loss: 0.05509922653436661
step: 90, loss: 0.040619637817144394
step: 100, loss: 0.006847796030342579
step: 110, loss: 0.010767086409032345
step: 120, loss: 0.031546447426080704
step: 130, loss: 0.006192827597260475
step: 140, loss: 0.032894067466259
step: 150, loss: 0.010400193743407726
step: 160, loss: 0.027257496491074562
step: 170, loss: 0.005100285168737173
step: 180, loss: 0.013375895097851753
step: 190, loss: 0.023482004180550575
step: 200, loss: 0.004758764058351517
step: 210, loss: 0.037162505090236664
step: 220, loss: 0.005605176091194153
step: 230, loss: 0.01760660856962204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776785714285714, 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: 0.002152330707758665
step: 10, loss: 0.016714364290237427
step: 20, loss: 0.003667321288958192
step: 30, loss: 0.0973047986626625
step: 40, loss: 0.04089047387242317
step: 50, loss: 0.005808192305266857
step: 60, loss: 0.004364493303000927
step: 70, loss: 0.006053212098777294
step: 80, loss: 0.007746092043817043
step: 90, loss: 0.0052159312181174755

```



```

step: 100, loss: 0.04098997265100479
step: 110, loss: 0.01161767728626728
step: 120, loss: 0.0009488392970524728
step: 130, loss: 0.09742624312639236
step: 140, loss: 0.009246145375072956
step: 150, loss: 0.007689904887229204
step: 160, loss: 0.023480437695980072
step: 170, loss: 0.01659817434847355
step: 180, loss: 0.006902368739247322
step: 190, loss: 0.0338541679084301
step: 200, loss: 0.013962113298475742
step: 210, loss: 0.006342145148664713
step: 220, loss: 0.029614578932523727
step: 230, loss: 0.059984318912029266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9797752808988766, f1=0.9738339021615473,
best_f1=0.9738339021615473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020447097718715668
step: 10, loss: 0.047659147530794144
step: 20, loss: 0.005924911703914404
step: 30, loss: 0.024110160768032074
step: 40, loss: 0.16072173416614532
step: 50, loss: 0.04391473904252052
step: 60, loss: 0.0012370948679745197
step: 70, loss: 0.008976190350949764
step: 80, loss: 0.008554428815841675
step: 90, loss: 0.07311351597309113
step: 100, loss: 0.0019566677510738373
step: 110, loss: 0.000724944518879056
step: 120, loss: 0.016757529228925705
step: 130, loss: 0.0032101869583129883
step: 140, loss: 0.0011112934444099665
step: 150, loss: 0.023560501635074615
step: 160, loss: 0.0028479553293436766
step: 170, loss: 0.026525339111685753
step: 180, loss: 0.1554606556892395
step: 190, loss: 0.012568368576467037
step: 200, loss: 0.005017217248678207
step: 210, loss: 0.0029737947043031454

```

```

step: 220, loss: 0.0004212723870296031
step: 230, loss: 0.0008643300388939679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9794988610478361,
best_f1=0.9794988610478361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004790644161403179
step: 10, loss: 0.0014521741541102529
step: 20, loss: 0.004302543122321367
step: 30, loss: 0.0018865809543058276
step: 40, loss: 0.002494210610166192
step: 50, loss: 0.001403767499141395
step: 60, loss: 0.002152565633878112
step: 70, loss: 0.006097923032939434
step: 80, loss: 0.09016671776771545
step: 90, loss: 0.02028871700167656
step: 100, loss: 0.0009413565858267248
step: 110, loss: 0.003815257688984275
step: 120, loss: 0.0177577193826437
step: 130, loss: 0.0011652050307020545
step: 140, loss: 0.002005005721002817
step: 150, loss: 0.20521670579910278
step: 160, loss: 0.0007081084186211228
step: 170, loss: 0.0011474663624539971
step: 180, loss: 0.026279713958501816
step: 190, loss: 0.17996928095817566
step: 200, loss: 0.0033068947959691286
step: 210, loss: 0.006118754390627146
step: 220, loss: 0.0005081529379822314
step: 230, loss: 0.004749342333525419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9844097995545658, f1=0.9832026875699889,
best_f1=0.9832026875699889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027138920268043876
step: 10, loss: 0.0008477148367092013
step: 20, loss: 0.00041580904508009553
step: 30, loss: 0.0003299191885162145
step: 40, loss: 0.00038560680695809424
step: 50, loss: 0.0004989439039491117
step: 60, loss: 0.0014089170144870877
step: 70, loss: 0.001614716136828065
step: 80, loss: 0.0023613388184458017
step: 90, loss: 0.0014763363869860768
step: 100, loss: 0.0017727743834257126
step: 110, loss: 0.002820944180712104
step: 120, loss: 0.0006101145409047604
step: 130, loss: 0.0003328098391648382
step: 140, loss: 0.0002608221548143774
step: 150, loss: 0.00018876248213928193
step: 160, loss: 0.030245784670114517
step: 170, loss: 0.00037454499397426844
step: 180, loss: 0.004144282545894384
step: 190, loss: 0.0005073304637335241
step: 200, loss: 0.0018254569731652737
step: 210, loss: 0.0008148018969222903
step: 220, loss: 0.027126645669341087
step: 230, loss: 0.001366323558613658
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9783845278725825,
best_f1=0.9832026875699889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00809458177536726
step: 10, loss: 0.0016036585438996553
step: 20, loss: 0.0005929336766712368
step: 30, loss: 0.00029409321723505855
step: 40, loss: 0.001485939952544868
step: 50, loss: 0.0029122550040483475
step: 60, loss: 0.000845845730509609
step: 70, loss: 0.0016104449750855565
step: 80, loss: 0.0004906732938252389
step: 90, loss: 0.0006583635113202035
step: 100, loss: 0.0002907077723648399
step: 110, loss: 0.0001747521891957149
step: 120, loss: 0.009233674965798855
step: 130, loss: 0.0018567581428214908

```

```

step: 140, loss: 0.0005441034445539117
step: 150, loss: 0.0479620024561882
step: 160, loss: 0.01450720988214016
step: 170, loss: 0.040409453213214874
step: 180, loss: 0.01275264099240303
step: 190, loss: 0.0002810113364830613
step: 200, loss: 0.000515724706929177
step: 210, loss: 0.010110745206475258
step: 220, loss: 0.0003494612465146929
step: 230, loss: 0.0015204299706965685
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9794988610478361,
best_f1=0.9832026875699889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0055420114658772945
step: 10, loss: 0.004294060170650482
step: 20, loss: 0.0029125558212399483
step: 30, loss: 0.0010206025326624513
step: 40, loss: 0.05777661129832268
step: 50, loss: 0.005475074518471956
step: 60, loss: 0.0012132767587900162
step: 70, loss: 0.00028315503732301295
step: 80, loss: 0.0005312474677339196
step: 90, loss: 0.0005652709514833987
step: 100, loss: 0.000547023955732584
step: 110, loss: 0.002380655612796545
step: 120, loss: 0.003051287494599819
step: 130, loss: 0.005014416761696339
step: 140, loss: 0.00020756082085426897
step: 150, loss: 0.019332116469740868
step: 160, loss: 0.020231636241078377
step: 170, loss: 0.02074773609638214
step: 180, loss: 0.0003525339998304844
step: 190, loss: 0.04908544197678566
step: 200, loss: 0.0004190345644019544
step: 210, loss: 0.00446687825024128
step: 220, loss: 0.0004533965256996453
step: 230, loss: 0.0035707764327526093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987598647125141, f1=0.9783352337514253,
best_f1=0.9783352337514253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002573600213509053
step: 10, loss: 0.0013903023209422827
step: 20, loss: 0.08327046036720276
step: 30, loss: 0.001682836445979774
step: 40, loss: 0.0007551686139777303
step: 50, loss: 0.009401229210197926
step: 60, loss: 0.0007382556213997304
step: 70, loss: 0.04257446900010109
step: 80, loss: 0.00046671091695316136
step: 90, loss: 0.0021205099765211344
step: 100, loss: 0.0009114614222198725
step: 110, loss: 0.000693188514560461
step: 120, loss: 0.031597670167684555
step: 130, loss: 0.0005843207472935319
step: 140, loss: 0.00044814791181124747
step: 150, loss: 0.0002851156168617308
step: 160, loss: 0.0007563602412119508
step: 170, loss: 0.00018350835307501256
step: 180, loss: 0.0006385222659446299
step: 190, loss: 0.0005067355814389884
step: 200, loss: 0.008137628436088562
step: 210, loss: 0.005627814214676619
step: 220, loss: 0.00013604768901132047
step: 230, loss: 0.002149284118786454
/usr/local/lib/python3.7/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=0.9864559819413092, f1=0.9783352337514253,
best_f1=0.9783352337514253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001273827801924199
step: 10, loss: 0.00023139981203712523
step: 20, loss: 0.001868553226813674
step: 30, loss: 0.00040031905518844724
step: 40, loss: 0.05512388050556183
step: 50, loss: 0.0001807277149055153
step: 60, loss: 0.00031001417664811015
step: 70, loss: 0.0007796682766638696
step: 80, loss: 0.000279885163763538
step: 90, loss: 0.010879083536565304

```

```

step: 100, loss: 0.0015989852836355567
step: 110, loss: 0.013211410492658615
step: 120, loss: 0.0001914891181513667
step: 130, loss: 0.00043564639054238796
step: 140, loss: 0.0003300171229057014
step: 150, loss: 0.0006215020548552275
step: 160, loss: 0.00018622472998686135
step: 170, loss: 0.0003403180744498968
step: 180, loss: 0.0005879136151634157
step: 190, loss: 0.00041056700865738094
step: 200, loss: 0.00041929923463612795
step: 210, loss: 0.011375082656741142
step: 220, loss: 0.0007503998349420726
step: 230, loss: 0.00024407934688497335
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9830124575311437,
best_f1=0.9783352337514253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001723280583973974
step: 10, loss: 0.0003209428978152573
step: 20, loss: 0.0008906511357054114
step: 30, loss: 0.00020334176952019334
step: 40, loss: 4.7678269766038284e-05
step: 50, loss: 0.00010432599810883403
step: 60, loss: 0.000637872377410531
step: 70, loss: 0.00032060753437690437
step: 80, loss: 0.0010805133497342467
step: 90, loss: 0.23642584681510925
step: 100, loss: 0.0007414681022055447
step: 110, loss: 0.0010145603446289897
step: 120, loss: 0.0003857540141325444
step: 130, loss: 0.02692572958767414
step: 140, loss: 0.0026992661878466606
step: 150, loss: 0.00046024442417547107
step: 160, loss: 0.0009010357316583395
step: 170, loss: 0.0011812306474894285
step: 180, loss: 0.0002777024346869439
step: 190, loss: 0.0002297381142852828
step: 200, loss: 0.0006316750659607351
step: 210, loss: 0.00028761778958141804
step: 220, loss: 0.0027502854354679585
step: 230, loss: 0.00020814796152990311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876265466816648, f1=0.9829738933030647,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041214292286895216
step: 10, loss: 0.00021845547598786652
step: 20, loss: 0.0018886583857238293
step: 30, loss: 0.007489476818591356
step: 40, loss: 0.0003908583603333682
step: 50, loss: 0.002717578085139394
step: 60, loss: 0.0018413026118651032
step: 70, loss: 0.0005179151194170117
step: 80, loss: 0.0026095816865563393
step: 90, loss: 0.0007683178409934044
step: 100, loss: 0.00036136049311608076
step: 110, loss: 0.0003275425115134567
step: 120, loss: 0.0003779589314945042
step: 130, loss: 0.0004184233257547021
step: 140, loss: 0.0006460165604948997
step: 150, loss: 0.0005756882019340992
step: 160, loss: 0.0017694057896733284
step: 170, loss: 0.0006256061606109142
step: 180, loss: 0.0002783832314889878
step: 190, loss: 0.0006643019150942564
step: 200, loss: 0.0003554157156031579
step: 210, loss: 0.0010020221816375852
step: 220, loss: 0.10791399329900742
step: 230, loss: 0.0005266336956992745
/usr/local/lib/python3.7/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=0.9865168539325843, f1=0.9830124575311437,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035879897768609226
step: 10, loss: 0.0009050092194229364
step: 20, loss: 0.0014086966402828693
step: 30, loss: 0.0002546822652220726
step: 40, loss: 0.0020735696889460087
step: 50, loss: 0.0006590208504348993

```

```

step: 60, loss: 0.0010051095159724355
step: 70, loss: 0.0002949814952444285
step: 80, loss: 0.000226684904191643
step: 90, loss: 0.0002815049374476075
step: 100, loss: 0.0007169591262936592
step: 110, loss: 0.00047933575115166605
step: 120, loss: 0.0005141579313203692
step: 130, loss: 0.0004441005876287818
step: 140, loss: 0.0002433951012790203
step: 150, loss: 0.0002534220984671265
step: 160, loss: 0.0007423476199619472
step: 170, loss: 0.0005935552180744708
step: 180, loss: 0.01630738563835621
step: 190, loss: 0.0005383103853091598
step: 200, loss: 0.00015078263822942972
step: 210, loss: 0.0005188702489249408
step: 220, loss: 0.0004994620103389025
step: 230, loss: 0.00046683268737979233
/usr/local/lib/python3.7/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=0.9864253393665158, f1=0.9794988610478361,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003853513044305146
step: 10, loss: 0.0002628158254083246
step: 20, loss: 0.0011409677099436522
step: 30, loss: 0.0007960381917655468
step: 40, loss: 0.0007264256710186601
step: 50, loss: 0.0003622094518505037
step: 60, loss: 0.0005225080531090498
step: 70, loss: 0.0004783083568327129
step: 80, loss: 0.0003029700310435146
step: 90, loss: 0.0007760459557175636
step: 100, loss: 0.0004671996575780213
step: 110, loss: 0.0007373301195912063
step: 120, loss: 0.0001564008998684585
step: 130, loss: 0.0006820109556429088
step: 140, loss: 0.0004387794469948858
step: 150, loss: 0.00022741440625395626
step: 160, loss: 0.00036883799475617707
step: 170, loss: 0.000491189886815846
step: 180, loss: 0.0004460293857846409
step: 190, loss: 0.0003338755923323333
step: 200, loss: 0.0004727420164272189
step: 210, loss: 0.0004947674460709095

```



```

step: 220, loss: 0.00037974922452121973
step: 230, loss: 0.020364901050925255
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9842342342342343,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018007151084020734
step: 10, loss: 0.00041233960655517876
step: 20, loss: 0.0005734785809181631
step: 30, loss: 0.0005054065841250122
step: 40, loss: 0.000254189275437966
step: 50, loss: 0.00040781660936772823
step: 60, loss: 0.0582381971180439
step: 70, loss: 0.001298578572459519
step: 80, loss: 0.0003895305562764406
step: 90, loss: 0.0003652451268862933
step: 100, loss: 0.00023183017037808895
step: 110, loss: 0.01875861920416355
step: 120, loss: 0.10117276757955551
step: 130, loss: 0.0002983348967973143
step: 140, loss: 0.0004709142376668751
step: 150, loss: 0.0012828484177589417
step: 160, loss: 0.0005664985510520637
step: 170, loss: 0.00021603642380796373
step: 180, loss: 0.0003899784933310002
step: 190, loss: 0.0006528373924084008
step: 200, loss: 0.0003658602072391659
step: 210, loss: 0.053606487810611725
step: 220, loss: 0.00035504333209246397
step: 230, loss: 0.0004274699022062123
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9842342342342343,
best_f1=0.9829738933030647
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                  : True
opt_level                : O2
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:16, 148.78it/s]
load_f1 = 0.9898762654668166
real_f1 = 0.9887387387387387
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 138.42it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.605427622795105

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

step: 10, loss: 0.4111347496509552
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.3536054193973541
step: 30, loss: 0.3180110454559326
step: 40, loss: 0.2886110544204712
step: 50, loss: 0.23285779356956482
step: 60, loss: 0.28456777334213257
step: 70, loss: 0.32895323634147644
step: 80, loss: 0.15355271100997925
step: 90, loss: 0.11921482533216476
step: 100, loss: 0.14362362027168274
step: 110, loss: 0.18808062374591827
step: 120, loss: 0.14982345700263977
step: 130, loss: 0.3590676486492157
step: 140, loss: 0.12211267650127411
step: 150, loss: 0.14746719598770142
step: 160, loss: 0.10761215537786484
step: 170, loss: 0.07251731306314468
step: 180, loss: 0.13572317361831665
step: 190, loss: 0.052894316613674164
step: 200, loss: 0.06348263472318649
step: 210, loss: 0.0525643527507782
step: 220, loss: 0.05781782791018486
step: 230, loss: 0.24743786454200745
step: 240, loss: 0.04115162789821625
step: 250, loss: 0.10684684664011002
step: 260, loss: 0.3341488540172577
step: 270, loss: 0.25711116194725037
step: 280, loss: 0.06603635102510452
step: 290, loss: 0.06632532179355621
step: 300, loss: 0.09252121299505234
step: 310, loss: 0.3359231650829315
step: 320, loss: 0.07602440565824509
step: 330, loss: 0.026114560663700104
step: 340, loss: 0.5348663926124573
step: 350, loss: 0.161882683634758
step: 360, loss: 0.06901216506958008
step: 370, loss: 0.010603583417832851
step: 380, loss: 0.2947230339050293
step: 390, loss: 0.04720043018460274
step: 400, loss: 0.05801163613796234
step: 410, loss: 0.3438384532928467
step: 420, loss: 0.05983962118625641
step: 430, loss: 0.0547533817589283
step: 440, loss: 0.0250373687595129
step: 450, loss: 0.08684232831001282
step: 460, loss: 0.021676668897271156
step: 470, loss: 0.04422910138964653

```

step: 480, loss: 0.1323043704032898
step: 490, loss: 0.1544944941997528
step: 500, loss: 0.03585585951805115
step: 510, loss: 0.026291515678167343
step: 520, loss: 0.04822354391217232
step: 530, loss: 0.1933268904685974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9282777523983554, f1=0.928635953026197,
best_f1=0.928635953026197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051302455365657806
step: 10, loss: 0.048655204474925995
step: 20, loss: 0.023255636915564537
step: 30, loss: 0.05585798993706703
step: 40, loss: 0.09201041609048843
step: 50, loss: 0.021238898858428
step: 60, loss: 0.04209292680025101
step: 70, loss: 0.056902337819337845
step: 80, loss: 0.034963496029376984
step: 90, loss: 0.10698071867227554
step: 100, loss: 0.125738725066185
step: 110, loss: 0.03828500956296921
step: 120, loss: 0.02230689488351345
step: 130, loss: 0.01412675715982914
step: 140, loss: 0.1146598607301712
step: 150, loss: 0.007210180629044771
step: 160, loss: 0.07335887104272842
step: 170, loss: 0.08660247921943665
step: 180, loss: 0.12934783101081848
step: 190, loss: 0.029249070212244987
step: 200, loss: 0.04245048388838768
step: 210, loss: 0.02071031928062439
step: 220, loss: 0.0012863441370427608
step: 230, loss: 0.06720717251300812
step: 240, loss: 0.05113938823342323
step: 250, loss: 0.02805960364639759
step: 260, loss: 0.07589193433523178
step: 270, loss: 0.12862886488437653
step: 280, loss: 0.07954108715057373
step: 290, loss: 0.05971790850162506

```

```

step: 300, loss: 0.1124582588672638
step: 310, loss: 0.021977009251713753
step: 320, loss: 0.0519854873418808
step: 330, loss: 0.14469438791275024
step: 340, loss: 0.05009184405207634
step: 350, loss: 0.003680651308968663
step: 360, loss: 0.20352791249752045
step: 370, loss: 0.07231482118368149
step: 380, loss: 0.15527015924453735
step: 390, loss: 0.016503388062119484
step: 400, loss: 0.07785150408744812
step: 410, loss: 0.04913782328367233
step: 420, loss: 0.026604054495692253
step: 430, loss: 0.14778169989585876
step: 440, loss: 0.02205786108970642
step: 450, loss: 0.022337473928928375
step: 460, loss: 0.11616884917020798
step: 470, loss: 0.15620683133602142
step: 480, loss: 0.0993831679224968
step: 490, loss: 0.10639883577823639
step: 500, loss: 0.03213227540254593
step: 510, loss: 0.04552767798304558
step: 520, loss: 0.3324168920516968
step: 530, loss: 0.04623844102025032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9427107591988821, f1=0.9399815327793167,
best_f1=0.9399815327793167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11543735861778259
step: 10, loss: 0.0602964423596859
step: 20, loss: 0.022910257801413536
step: 30, loss: 0.060348622500896454
step: 40, loss: 0.04303192347288132
step: 50, loss: 0.008303855545818806
step: 60, loss: 0.1865336298942566
step: 70, loss: 0.14293834567070007
step: 80, loss: 0.05399511009454727
step: 90, loss: 0.016119102016091347
step: 100, loss: 0.037569720298051834
step: 110, loss: 0.07154908031225204

```

```

step: 120, loss: 0.12968666851520538
step: 130, loss: 0.07089335471391678
step: 140, loss: 0.0096795205026865
step: 150, loss: 0.05922357365489006
step: 160, loss: 0.017367631196975708
step: 170, loss: 0.00206582248210907
step: 180, loss: 0.08788260817527771
step: 190, loss: 0.0021340991370379925
step: 200, loss: 0.020551927387714386
step: 210, loss: 0.04046421870589256
step: 220, loss: 0.10553435236215591
step: 230, loss: 0.03135930374264717
step: 240, loss: 0.08773521333932877
step: 250, loss: 0.07840297371149063
step: 260, loss: 0.017388256266713142
step: 270, loss: 0.004898421466350555
step: 280, loss: 0.0498928464949131
step: 290, loss: 0.007401874754577875
step: 300, loss: 0.21058806777000427
step: 310, loss: 0.03573589026927948
step: 320, loss: 0.041913364082574844
step: 330, loss: 0.006906920112669468
step: 340, loss: 0.007991495542228222
step: 350, loss: 0.021439163014292717
step: 360, loss: 0.0198814757168293
step: 370, loss: 0.034230440855026245
step: 380, loss: 0.010773113928735256
step: 390, loss: 0.016680320724844933
step: 400, loss: 0.03963698074221611
step: 410, loss: 0.1571054309606552
step: 420, loss: 0.008179902099072933
step: 430, loss: 0.019810881465673447
step: 440, loss: 0.21875372529029846
step: 450, loss: 0.012008097022771835
step: 460, loss: 0.1575096696615219
step: 470, loss: 0.003363661700859666
step: 480, loss: 0.07903403788805008
step: 490, loss: 0.014030409045517445
step: 500, loss: 0.020483246073126793
step: 510, loss: 0.0270547978579998
step: 520, loss: 0.013936067000031471
step: 530, loss: 0.014271173626184464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9453345900094251, f1=0.9427230046948356,
best_f1=0.9427230046948356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007338487543165684
step: 10, loss: 0.0011291225673630834
step: 20, loss: 0.29861390590667725
step: 30, loss: 0.009195908904075623
step: 40, loss: 0.01296486146748066
step: 50, loss: 0.11508522182703018
step: 60, loss: 0.02935813181102276
step: 70, loss: 0.05148676410317421
step: 80, loss: 0.047957394272089005
step: 90, loss: 0.04647814482450485
step: 100, loss: 0.04881986230611801
step: 110, loss: 0.05266382172703743
step: 120, loss: 0.0010491483844816685
step: 130, loss: 0.028611311689019203
step: 140, loss: 0.0465410053730011
step: 150, loss: 0.02124149724841118
step: 160, loss: 0.006745137739926577
step: 170, loss: 0.01850482076406479
step: 180, loss: 0.047547709196805954
step: 190, loss: 0.09380281716585159
step: 200, loss: 0.02184445597231388
step: 210, loss: 0.019653454422950745
step: 220, loss: 0.010998987592756748
step: 230, loss: 0.007876538671553135
step: 240, loss: 0.009420095942914486
step: 250, loss: 0.04793306812644005
step: 260, loss: 0.004815032705664635
step: 270, loss: 0.23180679976940155
step: 280, loss: 0.004286123439669609
step: 290, loss: 0.04955506697297096
step: 300, loss: 0.004020674619823694
step: 310, loss: 0.0037508418317884207
step: 320, loss: 0.4569686949253082
step: 330, loss: 0.0341792106628418
step: 340, loss: 0.0018024055752903223
step: 350, loss: 0.34015801548957825
step: 360, loss: 0.020802803337574005
step: 370, loss: 0.0016852435655891895
step: 380, loss: 0.0035139259416610003
step: 390, loss: 0.0002698109019547701
step: 400, loss: 0.05103446543216705
step: 410, loss: 0.025766370818018913
```



```
step: 420, loss: 0.013213477097451687
step: 430, loss: 0.021498702466487885
step: 440, loss: 0.012779180891811848
step: 450, loss: 0.014900612644851208
step: 460, loss: 0.030258623883128166
step: 470, loss: 0.002807522425428033
step: 480, loss: 0.002209589583799243
step: 490, loss: 0.024967603385448456
step: 500, loss: 0.0845707580447197
step: 510, loss: 0.07540683448314667
step: 520, loss: 0.013435822911560535
step: 530, loss: 0.04698776826262474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9452181987000929, f1=0.943344081068632,
best_f1=0.9427230046948356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03890848904848099
step: 10, loss: 0.004168728366494179
step: 20, loss: 0.004848246928304434
step: 30, loss: 0.001160487299785018
step: 40, loss: 0.00722834886983037
step: 50, loss: 0.07087526470422745
step: 60, loss: 0.018404340371489525
step: 70, loss: 0.007285307627171278
step: 80, loss: 0.023976173251867294
step: 90, loss: 0.09382876753807068
step: 100, loss: 0.0036967205815017223
step: 110, loss: 0.001850933418609202
step: 120, loss: 0.21643371880054474
step: 130, loss: 0.006522465031594038
step: 140, loss: 0.021499378606677055
step: 150, loss: 0.015097741968929768
step: 160, loss: 0.04779374226927757
step: 170, loss: 0.06493185460567474
step: 180, loss: 0.009733586572110653
step: 190, loss: 0.004719245247542858
step: 200, loss: 0.0024725806433707476
step: 210, loss: 0.0008129046182148159
step: 220, loss: 0.020887218415737152
step: 230, loss: 0.004710255656391382
step: 240, loss: 0.11042819172143936
step: 250, loss: 0.0675477534532547
step: 260, loss: 0.00036944547900930047
step: 270, loss: 0.00032750697573646903
```

```

step: 280, loss: 0.011925204657018185
step: 290, loss: 0.008653799071907997
step: 300, loss: 0.06477467715740204
step: 310, loss: 0.032623693346977234
step: 320, loss: 0.06182553619146347
step: 330, loss: 0.031862977892160416
step: 340, loss: 0.003832968883216381
step: 350, loss: 0.0004488248669076711
step: 360, loss: 0.00015934866678435355
step: 370, loss: 0.00014870309678371996
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 380, loss: 0.000491489889100194
step: 390, loss: 0.002788235666230321
step: 400, loss: 0.021277347579598427
step: 410, loss: 0.018503880128264427
step: 420, loss: 0.19139288365840912
step: 430, loss: 0.02096761390566826
step: 440, loss: 0.008880830369889736
step: 450, loss: 0.04034651070833206
step: 460, loss: 0.06484061479568481
step: 470, loss: 0.010292739607393742
step: 480, loss: 0.18028606474399567
step: 490, loss: 0.03015141747891903
step: 500, loss: 0.005402942653745413
step: 510, loss: 0.002754592103883624
step: 520, loss: 0.007434967905282974
step: 530, loss: 0.0017471034079790115
/usr/local/lib/python3.7/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=0.9427516158818098, f1=0.9372114496768237,
best_f1=0.9427230046948356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025713060051202774
step: 10, loss: 0.0006598843028768897
step: 20, loss: 0.005812221206724644
step: 30, loss: 0.001215625205077231
step: 40, loss: 0.004354786593466997
step: 50, loss: 0.09978681802749634
step: 60, loss: 0.009057041257619858
step: 70, loss: 0.04077880084514618
step: 80, loss: 0.0008299426408484578
step: 90, loss: 0.00825149193406105
step: 100, loss: 0.008659484796226025
step: 110, loss: 0.002135508693754673
step: 120, loss: 0.013286917470395565

```

```

step: 130, loss: 0.0002007904404308647
step: 140, loss: 0.006433961912989616
step: 150, loss: 0.00012418735423125327
step: 160, loss: 0.006413049530237913
step: 170, loss: 0.00019909994443878531
step: 180, loss: 0.020350221544504166
step: 190, loss: 0.048013247549533844
step: 200, loss: 0.010855543427169323
step: 210, loss: 0.008921013213694096
step: 220, loss: 0.02085774950683117
step: 230, loss: 0.006048617418855429
step: 240, loss: 0.05709947645664215
step: 250, loss: 0.24758659303188324
step: 260, loss: 0.011553283780813217
step: 270, loss: 0.013033510185778141
step: 280, loss: 0.015350133180618286
step: 290, loss: 0.026143336668610573
step: 300, loss: 0.007263100240379572
step: 310, loss: 0.09317942708730698
step: 320, loss: 0.005371949169784784
step: 330, loss: 0.008267831988632679
step: 340, loss: 0.0006370143964886665
step: 350, loss: 0.002712891437113285
step: 360, loss: 0.002246685622930527
step: 370, loss: 0.0227925144135952
step: 380, loss: 0.00973455049097538
step: 390, loss: 0.0004118474025744945
step: 400, loss: 0.025206129997968674
step: 410, loss: 0.0013622973347082734
step: 420, loss: 0.032977838069200516
step: 430, loss: 0.008952326141297817
step: 440, loss: 0.0007398242014460266
step: 450, loss: 0.1537775993347168
step: 460, loss: 0.033322058618068695
step: 470, loss: 0.0008555124513804913
step: 480, loss: 0.0028534969314932823
step: 490, loss: 0.005717015825212002
step: 500, loss: 0.002062366809695959
step: 510, loss: 0.24309447407722473
step: 520, loss: 0.004363050684332848
step: 530, loss: 0.017944909632205963
/usr/local/lib/python3.7/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=0.9390746678882271, f1=0.9413382218148486,
best_f1=0.9427230046948356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03417954221367836
step: 10, loss: 0.04239274561405182
step: 20, loss: 0.003997533582150936
step: 30, loss: 0.005787854082882404
step: 40, loss: 0.009644095785915852
step: 50, loss: 0.07287555932998657
step: 60, loss: 0.0015678575728088617
step: 70, loss: 0.025173509493470192
step: 80, loss: 0.004143375903367996
step: 90, loss: 0.0006064297631382942
step: 100, loss: 0.0067687369883060455
step: 110, loss: 0.007877035066485405
step: 120, loss: 0.0687251016497612
step: 130, loss: 0.0048208353109657764
step: 140, loss: 0.003119503380730748
step: 150, loss: 0.0018131956458091736
step: 160, loss: 0.0002019809471676126
step: 170, loss: 0.0035240775905549526
step: 180, loss: 0.0006209706189110875
step: 190, loss: 0.09645164012908936
step: 200, loss: 0.00037749207695014775
step: 210, loss: 0.041888028383255005
step: 220, loss: 0.10075800120830536
step: 230, loss: 0.00012310770398471504
step: 240, loss: 0.0003875130496453494
step: 250, loss: 0.0053026145324110985
step: 260, loss: 0.005410270299762487
step: 270, loss: 0.00923878699541092
step: 280, loss: 0.01232834067195654
step: 290, loss: 0.02077397145330906
step: 300, loss: 0.02351749688386917
step: 310, loss: 0.04949110001325607
step: 320, loss: 0.01919226534664631
step: 330, loss: 0.0013478173641487956
step: 340, loss: 0.0015196406748145819
step: 350, loss: 0.0012423018924891949
step: 360, loss: 0.0356648825109005
step: 370, loss: 0.0014588430058211088
step: 380, loss: 0.06196105107665062
step: 390, loss: 0.00526255089789629
step: 400, loss: 0.03004104644060135
step: 410, loss: 0.04524668678641319
step: 420, loss: 0.006490444298833609
step: 430, loss: 0.003614645916968584
step: 440, loss: 0.003896595910191536
step: 450, loss: 0.022510182112455368
step: 460, loss: 0.04792262613773346
```

```

step: 470, loss: 0.23715227842330933
step: 480, loss: 0.010446043685078621
step: 490, loss: 0.12626951932907104
step: 500, loss: 0.014442622661590576
step: 510, loss: 0.0035537993535399437
step: 520, loss: 0.001217709039337933
step: 530, loss: 0.0016899914480745792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9493554327808472, f1=0.9437070938215102,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00044580810936167836
step: 10, loss: 0.014469591900706291
step: 20, loss: 0.004503702744841576
step: 30, loss: 0.006887688301503658
step: 40, loss: 0.00011734855797840282
step: 50, loss: 0.014528095722198486
step: 60, loss: 0.0035564256832003593
step: 70, loss: 0.005499153397977352
step: 80, loss: 0.0008474360220134258
step: 90, loss: 0.00045297452015802264
step: 100, loss: 0.00011624764738371596
step: 110, loss: 0.00012849686027038842
step: 120, loss: 0.0014498346718028188
step: 130, loss: 0.0005463552661240101
step: 140, loss: 0.010593564249575138
step: 150, loss: 0.003107499796897173
step: 160, loss: 0.0004598013765644282
step: 170, loss: 0.004252840764820576
step: 180, loss: 0.018950704485177994
step: 190, loss: 0.004878249019384384
step: 200, loss: 0.005075821653008461
step: 210, loss: 0.04384990409016609
step: 220, loss: 0.0064105321653187275
step: 230, loss: 0.07720603048801422
step: 240, loss: 0.022157249972224236
step: 250, loss: 0.002578399144113064
step: 260, loss: 0.00266028824262321
step: 270, loss: 0.012629385106265545
step: 280, loss: 0.00023050520394463092

```

```

step: 290, loss: 0.00225139781832695
step: 300, loss: 0.00012349779717624187
step: 310, loss: 0.000506330921780318
step: 320, loss: 0.002735280664637685
step: 330, loss: 0.0004807534860447049
step: 340, loss: 0.0511753149330616
step: 350, loss: 0.006549629382789135
step: 360, loss: 0.0017007580026984215
step: 370, loss: 0.058446004986763
step: 380, loss: 6.544608913827688e-05
step: 390, loss: 0.10758928209543228
step: 400, loss: 0.0009184486116282642
step: 410, loss: 0.0022174601908773184
step: 420, loss: 0.0013420522445812821
step: 430, loss: 0.007477736100554466
step: 440, loss: 0.1248430609703064
step: 450, loss: 0.0013888846151530743
step: 460, loss: 0.007378846872597933
step: 470, loss: 0.01879078336060047
step: 480, loss: 0.0037731649354100227
step: 490, loss: 0.01549113355576992
step: 500, loss: 0.0015597864985466003
step: 510, loss: 0.044032178819179535
step: 520, loss: 0.0012858838308602571
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 530, loss: 0.0014178594574332237
/usr/local/lib/python3.7/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=0.9460853258321612, f1=0.9440820130475303,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00406331242993474
step: 10, loss: 0.0523357018828392
step: 20, loss: 0.0009205971146002412
step: 30, loss: 0.09142827987670898
step: 40, loss: 0.004493154585361481
step: 50, loss: 0.0012525771744549274
step: 60, loss: 0.0006055191624909639
step: 70, loss: 0.0020771909039467573
step: 80, loss: 0.013836700469255447
step: 90, loss: 0.0555749349296093
step: 100, loss: 0.005000447388738394
step: 110, loss: 0.004980797413736582
step: 120, loss: 0.010421494953334332
step: 130, loss: 0.002092648996040225

```

```

step: 140, loss: 0.012894740328192711
step: 150, loss: 0.0015933988615870476
step: 160, loss: 0.0013968036510050297
step: 170, loss: 0.0012765424326062202
step: 180, loss: 0.012493271380662918
step: 190, loss: 0.0007151992758736014
step: 200, loss: 0.00011496648949105293
step: 210, loss: 0.012269494123756886
step: 220, loss: 0.00024356390349566936
step: 230, loss: 0.0020240580197423697
step: 240, loss: 0.00022733015066478401
step: 250, loss: 0.0009169212426058948
step: 260, loss: 0.0023068636655807495
step: 270, loss: 0.009195471182465553
step: 280, loss: 0.022435206919908524
step: 290, loss: 0.0008677777368575335
step: 300, loss: 0.000547205563634634
step: 310, loss: 0.000435658119386062
step: 320, loss: 0.000892406387720257
step: 330, loss: 0.0016563410172238946
step: 340, loss: 0.07483293116092682
step: 350, loss: 0.010937116108834743
step: 360, loss: 0.08184877783060074
step: 370, loss: 0.0027236449532210827
step: 380, loss: 0.021591242402791977
step: 390, loss: 0.0002624003682285547
step: 400, loss: 0.0002417216746835038
step: 410, loss: 0.001764828572049737
step: 420, loss: 0.0005138781270943582
step: 430, loss: 0.04408719390630722
step: 440, loss: 0.0001425473456038162
step: 450, loss: 0.0014985488960519433
step: 460, loss: 0.0008317650062963367
step: 470, loss: 0.011918814852833748
step: 480, loss: 0.004745656158775091
step: 490, loss: 0.004180081654340029
step: 500, loss: 0.008711831644177437
step: 510, loss: 0.003428086405619979
step: 520, loss: 0.003526204265654087
step: 530, loss: 0.013810088858008385
/usr/local/lib/python3.7/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=0.943500229673863, f1=0.9459211732355638,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.0025255002547055483
step: 10, loss: 0.0008194977999664843
step: 20, loss: 0.0014908509328961372
step: 30, loss: 0.0010559775400906801
step: 40, loss: 0.0009833384538069367
step: 50, loss: 0.0034150180872529745
step: 60, loss: 0.006924901623278856
step: 70, loss: 0.00017898794612847269
step: 80, loss: 0.0011483489070087671
step: 90, loss: 9.775980288395658e-05
step: 100, loss: 0.002673278097063303
step: 110, loss: 0.006763840559870005
step: 120, loss: 0.00012247776612639427
step: 130, loss: 0.003592574503272772
step: 140, loss: 0.0007324857288040221
step: 150, loss: 9.594986477168277e-05
step: 160, loss: 0.049790266901254654
step: 170, loss: 0.02212333120405674
step: 180, loss: 0.017567476257681847
step: 190, loss: 0.0007235513185150921
step: 200, loss: 0.012490139342844486
step: 210, loss: 0.01776008866727352
step: 220, loss: 0.000517294043675065
step: 230, loss: 0.0004373917472548783
step: 240, loss: 0.001700124703347683
step: 250, loss: 0.002948337933048606
step: 260, loss: 0.0013141147792339325
step: 270, loss: 0.00017374982417095453
step: 280, loss: 0.016146713867783546
step: 290, loss: 0.0007265970925800502
step: 300, loss: 0.004437096416950226
step: 310, loss: 0.03474060446023941
step: 320, loss: 0.1021592915058136
step: 330, loss: 0.011986020021140575
step: 340, loss: 0.011000431142747402
step: 350, loss: 0.0012181381462141871
step: 360, loss: 0.0002111418143613264
step: 370, loss: 0.00426515145227313
step: 380, loss: 0.04788430035114288
step: 390, loss: 0.00043176536564715207
step: 400, loss: 0.0008596826810389757
step: 410, loss: 0.0005514483782462776
step: 420, loss: 0.00020795801538042724
step: 430, loss: 0.00036468260805122554
step: 440, loss: 0.013656344264745712
step: 450, loss: 0.0041485438123345375
step: 460, loss: 0.00018379731045570225
step: 470, loss: 0.01168761681765318


```

step: 480, loss: 0.0020477797370404005
step: 490, loss: 0.0011911464389413595
step: 500, loss: 0.0009186574025079608
step: 510, loss: 0.005581752397119999
step: 520, loss: 0.0012702887179329991
step: 530, loss: 0.04340782389044762
/usr/local/lib/python3.7/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=0.9439338235294118, f1=0.9422548120989918,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.690543962875381e-05
step: 10, loss: 0.0048721772618591785
step: 20, loss: 0.00014150208153296262
step: 30, loss: 0.0015955009730532765
step: 40, loss: 0.004985929466784
step: 50, loss: 3.200303763151169e-05
step: 60, loss: 0.0007782989414408803
step: 70, loss: 5.1429957238724455e-05
step: 80, loss: 0.0006446893094107509
step: 90, loss: 0.0005557136028073728
step: 100, loss: 0.00029861091752536595
step: 110, loss: 0.00016852625412866473
step: 120, loss: 0.0010026980889961123
step: 130, loss: 0.0033083544112741947
step: 140, loss: 0.0005610118969343603
step: 150, loss: 0.01614602655172348
step: 160, loss: 0.03342512622475624
step: 170, loss: 7.829828246030957e-05
step: 180, loss: 9.01694584172219e-05
step: 190, loss: 0.0004334042314440012
step: 200, loss: 0.0010035446612164378
step: 210, loss: 7.855003786971793e-05
step: 220, loss: 0.05032045394182205
step: 230, loss: 2.6362375137978233e-05
step: 240, loss: 0.0027040508575737476
step: 250, loss: 0.00046251717139966786
step: 260, loss: 0.0024229809641838074
step: 270, loss: 0.010778390802443027
step: 280, loss: 0.023836549371480942
step: 290, loss: 0.000967013300396502
step: 300, loss: 0.009966602548956871
step: 310, loss: 0.0026427526026964188
step: 320, loss: 0.0010095436591655016
step: 330, loss: 1.6998314094962552e-05

```

```

step: 340, loss: 0.0015579072060063481
step: 350, loss: 0.0018084325129166245
step: 360, loss: 0.00035217663389630616
step: 370, loss: 0.0019774625543504953
step: 380, loss: 0.0018231666181236506
step: 390, loss: 0.019591858610510826
step: 400, loss: 0.0026779579930007458
step: 410, loss: 0.0002657169825397432
step: 420, loss: 5.1638733566505834e-05
step: 430, loss: 0.027806328609585762
step: 440, loss: 0.0001627703895792365
step: 450, loss: 0.0011144231539219618
step: 460, loss: 0.045662641525268555
step: 470, loss: 0.00011198481661267579
step: 480, loss: 0.008024867624044418
step: 490, loss: 0.0009860798018053174
step: 500, loss: 0.00048229610547423363
step: 510, loss: 0.0013551110168918967
step: 520, loss: 0.00012947877985425293
step: 530, loss: 0.02154604159295559
/usr/local/lib/python3.7/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=0.9453526389537599, f1=0.9460465116279071,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000818756059743464
step: 10, loss: 0.00026983031420968473
step: 20, loss: 0.009660209529101849
step: 30, loss: 0.00032537724473513663
step: 40, loss: 0.0009627275285311043
step: 50, loss: 3.262691097916104e-05
step: 60, loss: 3.097748776781373e-05
step: 70, loss: 0.00014273091801442206
step: 80, loss: 0.055259134620428085
step: 90, loss: 0.013451233506202698
step: 100, loss: 0.3386433720588684
step: 110, loss: 0.0902843028306961
step: 120, loss: 0.0009233268210664392
step: 130, loss: 0.002060849918052554
step: 140, loss: 0.00023547446471638978
step: 150, loss: 5.228762165643275e-05
step: 160, loss: 0.0038055998738855124
step: 170, loss: 0.0005760032800026238
step: 180, loss: 0.000395656272303313
step: 190, loss: 0.0005436985520645976

```

```

step: 200, loss: 0.0034854498226195574
step: 210, loss: 0.000640267098788172
step: 220, loss: 0.0002314930607099086
step: 230, loss: 0.00010682023275876418
step: 240, loss: 0.0008261076873168349
step: 250, loss: 2.038821912719868e-05
step: 260, loss: 0.00014113058568909764
step: 270, loss: 0.3583052158355713
step: 280, loss: 9.706575656309724e-05
step: 290, loss: 0.002400388475507498
step: 300, loss: 0.0007841231999918818
step: 310, loss: 0.0012242471566423774
step: 320, loss: 0.08097943663597107
step: 330, loss: 0.0007315903203561902
step: 340, loss: 0.01047646813094616
step: 350, loss: 0.00012718680955003947
step: 360, loss: 0.0008931408519856632
step: 370, loss: 0.0018504724139347672
step: 380, loss: 0.0004244415904395282
step: 390, loss: 0.004015186335891485
step: 400, loss: 0.00010398222366347909
step: 410, loss: 0.0006389207555912435
step: 420, loss: 0.03824537247419357
step: 430, loss: 0.003753438824787736
step: 440, loss: 0.0054633584804832935
step: 450, loss: 0.010908582247793674
step: 460, loss: 0.0020440632943063974
step: 470, loss: 0.0013920952333137393
step: 480, loss: 0.0006431026267819107
step: 490, loss: 0.002094198251143098
step: 500, loss: 6.684681284241378e-05
step: 510, loss: 0.0034104110673069954
step: 520, loss: 0.0034458336886018515
step: 530, loss: 0.0016117044724524021
/usr/local/lib/python3.7/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=0.9454209065679925, f1=0.945337620578778,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003614892193581909
step: 10, loss: 0.004043797962367535
step: 20, loss: 9.116391447605565e-05
step: 30, loss: 0.00020657351706176996
step: 40, loss: 0.03569052368402481
step: 50, loss: 0.007048101630061865

```

step: 60, loss: 0.005916052032262087
step: 70, loss: 0.003097063861787319
step: 80, loss: 0.016735345125198364
step: 90, loss: 0.010258829221129417
step: 100, loss: 0.0003861792210955173
step: 110, loss: 0.006719681899994612
step: 120, loss: 0.00013334714458324015
step: 130, loss: 0.00209825299680233
step: 140, loss: 0.0003074311825912446
step: 150, loss: 0.002661022823303938
step: 160, loss: 0.001447186921723187
step: 170, loss: 0.0023714869748800993
step: 180, loss: 0.015188326127827168
step: 190, loss: 0.003613382112234831
step: 200, loss: 5.188227805774659e-05
step: 210, loss: 3.5766002838499844e-05
step: 220, loss: 0.001629730686545372
step: 230, loss: 0.01309810671955347
step: 240, loss: 0.009281081147491932
step: 250, loss: 0.0009434728417545557
step: 260, loss: 0.001401766436174512
step: 270, loss: 0.0010594993364065886
step: 280, loss: 0.0003502910549286753
step: 290, loss: 0.0006257395725697279
step: 300, loss: 0.0019107067491859198
step: 310, loss: 0.002343581523746252
step: 320, loss: 0.00028504422516562045
step: 330, loss: 0.001922102877870202
step: 340, loss: 0.0028917917516082525
step: 350, loss: 0.0009132064296863973
step: 360, loss: 0.008159332908689976
step: 370, loss: 0.002216402906924486
step: 380, loss: 0.0001633213832974434
step: 390, loss: 0.00023951819457579404
step: 400, loss: 0.0004856713640037924
step: 410, loss: 0.0007970579899847507
step: 420, loss: 0.0011366592952981591
step: 430, loss: 0.0007501919753849506
step: 440, loss: 0.0035200633574277163
step: 450, loss: 0.0005670630489476025
step: 460, loss: 0.017587924376130104
step: 470, loss: 0.0024345996789634228
step: 480, loss: 1.3030930858803913e-05
step: 490, loss: 0.0018462385050952435
step: 500, loss: 0.000672902911901474
step: 510, loss: 0.0008402427192777395
step: 520, loss: 0.00020315940491855145
step: 530, loss: 0.0009527496877126396

```

/usr/local/lib/python3.7/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=0.9468822170900694, f1=0.9442653155228006,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003164565423503518
step: 10, loss: 0.004577468615025282
step: 20, loss: 0.0002228360972367227
step: 30, loss: 0.0001514497707830742
step: 40, loss: 0.00012760484241880476
step: 50, loss: 0.0005459673702716827
step: 60, loss: 0.0002832703466992825
step: 70, loss: 4.82008617836982e-05
step: 80, loss: 0.00021021731663495302
step: 90, loss: 1.1168322998855729e-05
step: 100, loss: 0.000664238294120878
step: 110, loss: 0.0008528695907443762
step: 120, loss: 1.1376894690329209e-05
step: 130, loss: 0.00017774017760530114
step: 140, loss: 0.0033325699623674154
step: 150, loss: 1.944514224305749e-05
step: 160, loss: 0.0008500466356053948
step: 170, loss: 0.017899179831147194
step: 180, loss: 0.00013213089550845325
step: 190, loss: 2.3062279069563374e-05
step: 200, loss: 0.03108730912208557
step: 210, loss: 7.370815728791058e-05
step: 220, loss: 7.905646634753793e-05
step: 230, loss: 0.003320708405226469
step: 240, loss: 0.04202523082494736
step: 250, loss: 0.0007825850625522435
step: 260, loss: 3.603666482376866e-05
step: 270, loss: 0.0013243536232039332
step: 280, loss: 0.004388726782053709
step: 290, loss: 5.6632143241586164e-05
step: 300, loss: 0.00035626726457849145
step: 310, loss: 0.00022959178022574633
step: 320, loss: 4.869868644163944e-05
step: 330, loss: 0.000199960995814763
step: 340, loss: 0.00043685580021701753
step: 350, loss: 0.00020679701992776245
step: 360, loss: 0.0005464706919156015
step: 370, loss: 0.010467116720974445
step: 380, loss: 0.1092333197593689
step: 390, loss: 0.0009696815977804363

```

```

step: 400, loss: 0.0030063700396567583
step: 410, loss: 7.29054881958291e-05
step: 420, loss: 7.282858859980479e-05
step: 430, loss: 0.05497593805193901
step: 440, loss: 0.08364316821098328
step: 450, loss: 2.4220436898758635e-05
step: 460, loss: 0.00422176206484437
step: 470, loss: 0.00012243851961102337
step: 480, loss: 9.02074680197984e-05
step: 490, loss: 2.2134932805784047e-05
step: 500, loss: 0.0018917920533567667
step: 510, loss: 0.01347214262932539
step: 520, loss: 0.00013556079647969455
step: 530, loss: 0.0010074382880702615
/usr/local/lib/python3.7/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=0.9465437788018433, f1=0.9471750114836931,
best_f1=0.9437070938215102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014736002776771784
step: 10, loss: 0.0014147281181067228
step: 20, loss: 0.001290271757170558
step: 30, loss: 0.007156145293265581
step: 40, loss: 4.3895204726140946e-05
step: 50, loss: 0.00651507405564189
step: 60, loss: 0.004620752297341824
step: 70, loss: 6.244915130082518e-05
step: 80, loss: 0.00020669816876761615
step: 90, loss: 0.00020213061361573637
step: 100, loss: 2.777096233330667e-05
step: 110, loss: 0.02222294732928276
step: 120, loss: 5.097210669191554e-05
step: 130, loss: 0.0006746685248799622
step: 140, loss: 2.5729119442985393e-05
step: 150, loss: 0.00030385819263756275
step: 160, loss: 9.878297714749351e-05
step: 170, loss: 0.00013832429249305278
step: 180, loss: 0.003778154030442238
step: 190, loss: 0.0010575432097539306
step: 200, loss: 0.002199730835855007
step: 210, loss: 0.00017684900376480073
step: 220, loss: 0.0003641037328634411
step: 230, loss: 0.0006024862523190677
step: 240, loss: 0.003758138045668602
step: 250, loss: 0.001415903214365244

```

```

step: 260, loss: 0.00022706879826728255
step: 270, loss: 0.00013633280468638986
step: 280, loss: 0.0005368154379539192
step: 290, loss: 0.0004913238226436079
step: 300, loss: 0.00039854663191363215
step: 310, loss: 0.06061755493283272
step: 320, loss: 4.3784752051578835e-05
step: 330, loss: 8.455680654151365e-05
step: 340, loss: 0.0014087799936532974
step: 350, loss: 0.026668675243854523
step: 360, loss: 0.00012035293184453622
step: 370, loss: 0.0026289590168744326
step: 380, loss: 5.792501906398684e-05
step: 390, loss: 0.00297488272190094
step: 400, loss: 0.012106876820325851
step: 410, loss: 4.30717300332617e-05
step: 420, loss: 0.000236569409025833
step: 430, loss: 3.9741538785165176e-05
step: 440, loss: 0.0007937222835607827
step: 450, loss: 0.00047168118180707097
step: 460, loss: 5.715906809200533e-05
step: 470, loss: 1.4569363884220365e-05
step: 480, loss: 0.003927586600184441
step: 490, loss: 0.0004720157594420016
step: 500, loss: 0.00013816854334436357
step: 510, loss: 0.012948034331202507
step: 520, loss: 5.472189877764322e-05
step: 530, loss: 0.00021481048315763474
/usr/local/lib/python3.7/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=0.9457579972183587, f1=0.9494199535962878,
best_f1=0.9437070938215102
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:33, 171.85it/s]
load_f1 = 0.9450236966824644
real_f1 = 0.9419047619047619
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 145.41it/s]

```

4.1.4 Amazon-Google - Running the matcher

```

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

```



```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.538643479347229
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5021615028381348
step: 20, loss: 0.4476085901260376
step: 30, loss: 0.40128442645072937
step: 40, loss: 0.32623711228370667
step: 50, loss: 0.43407297134399414
step: 60, loss: 0.5084262490272522
step: 70, loss: 0.30423980951309204
step: 80, loss: 0.3808304965496063
step: 90, loss: 0.2377193719148636
step: 100, loss: 0.2311309576034546
step: 110, loss: 0.3061896562576294
step: 120, loss: 0.39354220032691956
step: 130, loss: 0.24938878417015076
step: 140, loss: 0.4638885259628296
step: 150, loss: 0.38745957612991333
step: 160, loss: 0.49852728843688965
step: 170, loss: 0.24730806052684784
step: 180, loss: 0.37303775548934937
step: 190, loss: 0.5937467813491821
step: 200, loss: 0.3101876974105835
step: 210, loss: 0.3939567506313324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.31185807656395886, f1=0.3480947476828012,
best_f1=0.3480947476828012
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3561165928840637
step: 10, loss: 0.15229305624961853
step: 20, loss: 0.3309703469276428
step: 30, loss: 0.4845881462097168
step: 40, loss: 0.4363008141517639
step: 50, loss: 0.22693242132663727
step: 60, loss: 0.33079826831817627
step: 70, loss: 0.25658679008483887
step: 80, loss: 0.25173935294151306
step: 90, loss: 0.23122450709342957
step: 100, loss: 0.6365219354629517

```

```

step: 110, loss: 0.3659083843231201
step: 120, loss: 0.20616765320301056
step: 130, loss: 0.18224278092384338
step: 140, loss: 0.2407030612230301
step: 150, loss: 0.3683202564716339
step: 160, loss: 0.13985909521579742
step: 170, loss: 0.2969287037849426
step: 180, loss: 0.27079859375953674
step: 190, loss: 0.33610963821411133
step: 200, loss: 0.12477336823940277
step: 210, loss: 0.2866644859313965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4750499001996008, f1=0.4528301886792453,
best_f1=0.4528301886792453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10563086718320847
step: 10, loss: 0.13878871500492096
step: 20, loss: 0.3099583387374878
step: 30, loss: 0.1213889941573143
step: 40, loss: 0.29072561860084534
step: 50, loss: 0.33763882517814636
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.24666035175323486
step: 70, loss: 0.2668120265007019
step: 80, loss: 0.19742262363433838
step: 90, loss: 0.18331661820411682
step: 100, loss: 0.2685087323188782
step: 110, loss: 0.1048516035079956
step: 120, loss: 0.17819201946258545
step: 130, loss: 0.38868457078933716
step: 140, loss: 0.2818925678730011
step: 150, loss: 0.22284753620624542
step: 160, loss: 0.1546260565519333
step: 170, loss: 0.2610618472099304
step: 180, loss: 0.12767158448696136
step: 190, loss: 0.06526508927345276
step: 200, loss: 0.17238734662532806
step: 210, loss: 0.26669013500213623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5294117647058824, f1=0.552808988764045,
best_f1=0.552808988764045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1533452272415161
step: 10, loss: 0.156441330909729
step: 20, loss: 0.20530840754508972
step: 30, loss: 0.06925883889198303
step: 40, loss: 0.06042968109250069
step: 50, loss: 0.1955358237028122
step: 60, loss: 0.2309178113937378
step: 70, loss: 0.13803884387016296
step: 80, loss: 0.07328818738460541
step: 90, loss: 0.12883742153644562
step: 100, loss: 0.24305981397628784
step: 110, loss: 0.4177044928073883
step: 120, loss: 0.27115583419799805
step: 130, loss: 0.29649069905281067
step: 140, loss: 0.48379620909690857
step: 150, loss: 0.22949199378490448
step: 160, loss: 0.357779324054718
step: 170, loss: 0.23183603584766388
step: 180, loss: 0.05560741201043129
step: 190, loss: 0.09964612126350403
step: 200, loss: 0.0816098228096962
step: 210, loss: 0.36803290247917175
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5495495495495496, f1=0.5159817351598174,
best_f1=0.5159817351598174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21611563861370087
step: 10, loss: 0.14564475417137146
step: 20, loss: 0.14936983585357666
step: 30, loss: 0.047100216150283813

```

```

step: 40, loss: 0.12058819085359573
step: 50, loss: 0.3234191834926605
step: 60, loss: 0.10527709126472473
step: 70, loss: 0.18835610151290894
step: 80, loss: 0.10605703294277191
step: 90, loss: 0.1638578474521637
step: 100, loss: 0.019470885396003723
step: 110, loss: 0.07580427825450897
step: 120, loss: 0.17941367626190186
step: 130, loss: 0.1163836270570755
step: 140, loss: 0.1294906735420227
step: 150, loss: 0.11777254194021225
step: 160, loss: 0.05676382780075073
step: 170, loss: 0.08044175803661346
step: 180, loss: 0.3997303545475006
step: 190, loss: 0.21860407292842865
step: 200, loss: 0.2298741191625595
step: 210, loss: 0.15511395037174225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5514950166112956, f1=0.5622895622895623,
best_f1=0.5622895622895623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040727101266384125
step: 10, loss: 0.18562230467796326
step: 20, loss: 0.11435827612876892
step: 30, loss: 0.18928292393684387
step: 40, loss: 0.10137119144201279
step: 50, loss: 0.07128068804740906
step: 60, loss: 0.21754764020442963
step: 70, loss: 0.16947060823440552
step: 80, loss: 0.08230747282505035
step: 90, loss: 0.034621402621269226
step: 100, loss: 0.1586347073316574
step: 110, loss: 0.2126205563545227
step: 120, loss: 0.044168055057525635
step: 130, loss: 0.09793004393577576
step: 140, loss: 0.050382234156131744
step: 150, loss: 0.08164083957672119
step: 160, loss: 0.07396254688501358
step: 170, loss: 0.08552026748657227

```

```

step: 180, loss: 0.11645577847957611
step: 190, loss: 0.045845143496990204
step: 200, loss: 0.13112998008728027
step: 210, loss: 0.049060966819524765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.576923076923077, f1=0.5833333333333334,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08506832271814346
step: 10, loss: 0.0208616703748703
step: 20, loss: 0.044269341975450516
step: 30, loss: 0.038894809782505035
step: 40, loss: 0.008749839849770069
step: 50, loss: 0.12745459377765656
step: 60, loss: 0.052325427532196045
step: 70, loss: 0.03948220610618591
step: 80, loss: 0.17707529664039612
step: 90, loss: 0.17678380012512207
step: 100, loss: 0.08999624848365784
step: 110, loss: 0.18372659385204315
step: 120, loss: 0.26386961340904236
step: 130, loss: 0.22639986872673035
step: 140, loss: 0.10443063080310822
step: 150, loss: 0.19312651455402374
step: 160, loss: 0.3417530357837677
step: 170, loss: 0.10540266335010529
step: 180, loss: 0.025183701887726784
step: 190, loss: 0.047123972326517105
step: 200, loss: 0.22950729727745056
step: 210, loss: 0.20182035863399506
/usr/local/lib/python3.7/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=0.5519848771266541, f1=0.5390624999999999,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08055198937654495
step: 10, loss: 0.038064830005168915

```

```
step: 20, loss: 0.016824159771203995
step: 30, loss: 0.018102647736668587
step: 40, loss: 0.11573062092065811
step: 50, loss: 0.04354773834347725
step: 60, loss: 0.0323767364025116
step: 70, loss: 0.2587686777114868
step: 80, loss: 0.11727732419967651
step: 90, loss: 0.1538926661014557
step: 100, loss: 0.2820439636707306
step: 110, loss: 0.034571900963783264
step: 120, loss: 0.13786455988883972
step: 130, loss: 0.07598751038312912
step: 140, loss: 0.026086188852787018
step: 150, loss: 0.19436955451965332
step: 160, loss: 0.1420975774526596
step: 170, loss: 0.045292895287275314
step: 180, loss: 0.0665680393576622
step: 190, loss: 0.13700541853904724
step: 200, loss: 0.07422225177288055
step: 210, loss: 0.22380666434764862
/usr/local/lib/python3.7/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=0.5573192239858906, f1=0.5545617173524151,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026959717273712158
step: 10, loss: 0.031572166830301285
step: 20, loss: 0.2725944519042969
step: 30, loss: 0.02582068368792534
step: 40, loss: 0.02128990739583969
step: 50, loss: 0.023337336257100105
step: 60, loss: 0.09175331890583038
step: 70, loss: 0.15318866074085236
step: 80, loss: 0.10533594340085983
step: 90, loss: 0.062223490327596664
step: 100, loss: 0.019953053444623947
step: 110, loss: 0.1755640208721161
step: 120, loss: 0.15944640338420868
step: 130, loss: 0.033336248248815536
step: 140, loss: 0.1250125765800476
step: 150, loss: 0.005701341200619936
step: 160, loss: 0.09403658658266068
step: 170, loss: 0.04716639220714569
step: 180, loss: 0.061353325843811035
step: 190, loss: 0.029874350875616074
```

```

step: 200, loss: 0.00951545313000679
step: 210, loss: 0.1883133500814438
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.577391304347826, f1=0.5743944636678201,
best_f1=0.5743944636678201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0461442805826664
step: 10, loss: 0.018693726509809494
step: 20, loss: 0.0986737310886383
step: 30, loss: 0.013285118155181408
step: 40, loss: 0.20883478224277496
step: 50, loss: 0.06343546509742737
step: 60, loss: 0.0352354571223259
step: 70, loss: 0.0901423767209053
step: 80, loss: 0.045582130551338196
step: 90, loss: 0.09933852404356003
step: 100, loss: 0.01666327938437462
step: 110, loss: 0.004125657957047224
step: 120, loss: 0.04725564271211624
step: 130, loss: 0.10810857266187668
step: 140, loss: 0.01900581084191799
step: 150, loss: 0.0425407849252224
step: 160, loss: 0.05506468936800957
step: 170, loss: 0.061895065009593964
step: 180, loss: 0.12607645988464355
step: 190, loss: 0.029345158487558365
step: 200, loss: 0.24057123064994812
step: 210, loss: 0.042916327714920044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5776173285198556, f1=0.5636363636363636,
best_f1=0.5636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.05659444257616997
step: 10, loss: 0.10663527995347977
step: 20, loss: 0.11896409094333649
step: 30, loss: 0.052468836307525635
step: 40, loss: 0.07352615147829056
step: 50, loss: 0.1700037568807602
step: 60, loss: 0.035628531128168106
step: 70, loss: 0.05912966653704643
step: 80, loss: 0.023515230044722557
step: 90, loss: 0.21267150342464447
step: 100, loss: 0.15781208872795105
step: 110, loss: 0.09278381615877151
step: 120, loss: 0.06800127774477005
step: 130, loss: 0.008233319036662579
step: 140, loss: 0.044185295701026917
step: 150, loss: 0.08366113901138306
step: 160, loss: 0.012383567169308662
step: 170, loss: 0.024178432300686836
step: 180, loss: 0.015630021691322327
step: 190, loss: 0.11502224206924438
step: 200, loss: 0.005126310512423515
step: 210, loss: 0.15676166117191315
/usr/local/lib/python3.7/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=0.5644859813084112, f1=0.5687022900763359,
best_f1=0.5636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0064929272048175335
step: 10, loss: 0.024979762732982635
step: 20, loss: 0.30763012170791626
step: 30, loss: 0.0032886324916034937
step: 40, loss: 0.004304781090468168
step: 50, loss: 0.010973693802952766
step: 60, loss: 0.0029624144081026316
step: 70, loss: 0.0595511868596077
step: 80, loss: 0.05708540230989456
step: 90, loss: 0.044103726744651794
step: 100, loss: 0.0032316658180207014
step: 110, loss: 0.28426867723464966
step: 120, loss: 0.008680516853928566
step: 130, loss: 0.17881950736045837
step: 140, loss: 0.06477747857570648
step: 150, loss: 0.007610783446580172
step: 160, loss: 0.04243692010641098
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

```

```

step: 170, loss: 0.11335094273090363
step: 180, loss: 0.016204621642827988
step: 190, loss: 0.12751998007297516
step: 200, loss: 0.01348910853266716
step: 210, loss: 0.036087214946746826
/usr/local/lib/python3.7/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=0.5662650602409638, f1=0.5685884691848907,
best_f1=0.5636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022048339247703552
step: 10, loss: 0.06388279050588608
step: 20, loss: 0.004726485814899206
step: 30, loss: 0.0048180557787418365
step: 40, loss: 0.007597280200570822
step: 50, loss: 0.06469854712486267
step: 60, loss: 0.022615674883127213
step: 70, loss: 0.0041962554678320885
step: 80, loss: 0.1709117740392685
step: 90, loss: 0.169192835688591
step: 100, loss: 0.04703215882182121
step: 110, loss: 0.021116379648447037
step: 120, loss: 0.010929719544947147
step: 130, loss: 0.013737057335674763
step: 140, loss: 0.020436104387044907
step: 150, loss: 0.010684476234018803
step: 160, loss: 0.015594201162457466
step: 170, loss: 0.04591802880167961
step: 180, loss: 0.045581188052892685
step: 190, loss: 0.02879912778735161
step: 200, loss: 0.09798067808151245
step: 210, loss: 0.011731826700270176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5784832451499118, f1=0.5601374570446735,
best_f1=0.5601374570446735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023561066016554832

```

```

step: 10, loss: 0.04282836243510246
step: 20, loss: 0.010838489048182964
step: 30, loss: 0.01042183954268694
step: 40, loss: 0.023593997582793236
step: 50, loss: 0.02212141454219818
step: 60, loss: 0.09005160629749298
step: 70, loss: 0.0038448290433734655
step: 80, loss: 0.07429873198270798
step: 90, loss: 0.031213970854878426
step: 100, loss: 0.008350012823939323
step: 110, loss: 0.0018733299802988768
step: 120, loss: 0.012933051213622093
step: 130, loss: 0.010068250820040703
step: 140, loss: 0.008150403387844563
step: 150, loss: 0.06932409852743149
step: 160, loss: 0.10382696241140366
step: 170, loss: 0.025889353826642036
step: 180, loss: 0.0033901468850672245
step: 190, loss: 0.04732269421219826
step: 200, loss: 0.006372748874127865
step: 210, loss: 0.011882795952260494
/usr/local/lib/python3.7/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=0.5645161290322581, f1=0.5877551020408163,
best_f1=0.5601374570446735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015669193118810654
step: 10, loss: 0.011120756156742573
step: 20, loss: 0.06217954680323601
step: 30, loss: 0.005931207910180092
step: 40, loss: 0.01423646230250597
step: 50, loss: 0.007057589013129473
step: 60, loss: 0.011440320871770382
step: 70, loss: 0.010440738871693611
step: 80, loss: 0.01553881075233221
step: 90, loss: 0.003972141537815332
step: 100, loss: 0.05179296433925629
step: 110, loss: 0.05569665879011154
step: 120, loss: 0.025677641853690147
step: 130, loss: 0.035271357744932175
step: 140, loss: 0.0029590344056487083
step: 150, loss: 0.06800524145364761
step: 160, loss: 0.2170429676771164
step: 170, loss: 0.02854899875819683
step: 180, loss: 0.014822997152805328

```

```

step: 190, loss: 0.05867817625403404
step: 200, loss: 0.005195945966988802
step: 210, loss: 0.11536506563425064
/usr/local/lib/python3.7/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=0.5770750988142292, f1=0.5823293172690762,
best_f1=0.5601374570446735
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 [00:10, 221.70it/s]
load_f1 = 0.5765765765765765
real_f1 = 0.5765765765765765
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 146.43it/s]

```

4.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.46305420994758606
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5056207776069641
step: 20, loss: 0.2534935474395752
step: 30, loss: 0.3781810700893402
step: 40, loss: 0.24309833347797394
step: 50, loss: 0.3213846981525421
step: 60, loss: 0.4451787769794464
step: 70, loss: 0.4538276791572571
step: 80, loss: 0.15776598453521729
step: 90, loss: 0.3474666178226471
step: 100, loss: 0.42364779114723206
step: 110, loss: 0.25525665283203125
step: 120, loss: 0.3384707570075989
step: 130, loss: 0.3018757402896881
step: 140, loss: 0.19486217200756073
step: 150, loss: 0.2965104281902313
step: 160, loss: 0.21978066861629486
step: 170, loss: 0.36781471967697144
step: 180, loss: 0.1656670719385147
step: 190, loss: 0.16408725082874298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4195658564567566
step: 10, loss: 0.30505070090293884
step: 20, loss: 0.6202569007873535
step: 30, loss: 0.24016988277435303
step: 40, loss: 0.5469989776611328
step: 50, loss: 0.3142871558666229
step: 60, loss: 0.44843798875808716
step: 70, loss: 0.3648032248020172
step: 80, loss: 0.16415809094905853
step: 90, loss: 0.3093292713165283
step: 100, loss: 0.2621834874153137
step: 110, loss: 0.3712897002696991
step: 120, loss: 0.23907731473445892
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.541818380355835
step: 140, loss: 0.31032833456993103
step: 150, loss: 0.3193493187427521
step: 160, loss: 0.3143823742866516
step: 170, loss: 0.23877876996994019
step: 180, loss: 0.17228060960769653
step: 190, loss: 0.24048446118831635
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39211225509643555
step: 10, loss: 0.37998276948928833
step: 20, loss: 0.4643968641757965
step: 30, loss: 0.34167030453681946
step: 40, loss: 0.0826711654663086
step: 50, loss: 0.3705313503742218
step: 60, loss: 0.1696474552154541
step: 70, loss: 0.3873235881328583
step: 80, loss: 0.3134828507900238

```

```

step: 90, loss: 0.3757866621017456
step: 100, loss: 0.543746292591095
step: 110, loss: 0.6706368923187256
step: 120, loss: 0.3837617039680481
step: 130, loss: 0.15571753680706024
step: 140, loss: 0.3802715837955475
step: 150, loss: 0.3092266917228699
step: 160, loss: 0.6077232360839844
step: 170, loss: 0.4525311589241028
step: 180, loss: 0.37706896662712097
step: 190, loss: 0.16762974858283997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23889631032943726
step: 10, loss: 0.23558789491653442
step: 20, loss: 0.3072218596935272
step: 30, loss: 0.23847679793834686
step: 40, loss: 0.5711105465888977
step: 50, loss: 0.23520679771900177
step: 60, loss: 0.3931310772895813
step: 70, loss: 0.33041101694107056
step: 80, loss: 0.2388388067483902
step: 90, loss: 0.17496874928474426
step: 100, loss: 0.3062538802623749
step: 110, loss: 0.3711698055267334
step: 120, loss: 0.23673215508460999
step: 130, loss: 0.46654799580574036
step: 140, loss: 0.31169816851615906
step: 150, loss: 0.24988004565238953
step: 160, loss: 0.30834558606147766
step: 170, loss: 0.4485637843608856
step: 180, loss: 0.38023969531059265
step: 190, loss: 0.16094639897346497
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39278122782707214

```



```

step: 10, loss: 0.39722543954849243
step: 20, loss: 0.17034846544265747
step: 30, loss: 0.1155373826622963
step: 40, loss: 0.30483588576316833
step: 50, loss: 0.5165606141090393
step: 60, loss: 0.23379208147525787
step: 70, loss: 0.39203715324401855
step: 80, loss: 0.38270607590675354
step: 90, loss: 0.30288538336753845
step: 100, loss: 0.45129045844078064
step: 110, loss: 0.4168691039085388
step: 120, loss: 0.22904232144355774
step: 130, loss: 0.552129328250885
step: 140, loss: 0.3766102194786072
step: 150, loss: 0.3000763952732086
step: 160, loss: 0.16482357680797577
step: 170, loss: 0.36924558877944946
step: 180, loss: 0.23707903921604156
step: 190, loss: 0.3125121593475342
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3082890808582306
step: 10, loss: 0.2418086677789688
step: 20, loss: 0.3141419589519501
step: 30, loss: 0.4704611301422119
step: 40, loss: 0.2359425574541092
step: 50, loss: 0.31022733449935913
step: 60, loss: 0.4473911225795746
step: 70, loss: 0.31311097741127014
step: 80, loss: 0.31078723073005676
step: 90, loss: 0.23257426917552948
step: 100, loss: 0.48473843932151794
step: 110, loss: 0.23451019823551178
step: 120, loss: 0.45293453335762024
step: 130, loss: 0.5191466808319092
step: 140, loss: 0.18436384201049805
step: 150, loss: 0.37553977966308594
step: 160, loss: 0.3773964047431946
step: 170, loss: 0.3889712393283844
step: 180, loss: 0.17929750680923462
step: 190, loss: 0.31267249584198
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3120790421962738
step: 10, loss: 0.31261810660362244
step: 20, loss: 0.31375619769096375
step: 30, loss: 0.1919134259223938
step: 40, loss: 0.3169578015804291
step: 50, loss: 0.08215031772851944
step: 60, loss: 0.15031251311302185
step: 70, loss: 0.15941083431243896
step: 80, loss: 0.23988018929958344
step: 90, loss: 0.2339056134223938
step: 100, loss: 0.5896197557449341
step: 110, loss: 0.44378525018692017
step: 120, loss: 0.38921672105789185
step: 130, loss: 0.3124792277812958
step: 140, loss: 0.24356772005558014
step: 150, loss: 0.30172550678253174
step: 160, loss: 0.3881551921367645
step: 170, loss: 0.30197450518608093
step: 180, loss: 0.23801808059215546
step: 190, loss: 0.31366997957229614
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23416224122047424
step: 10, loss: 0.3879377543926239
step: 20, loss: 0.2460479587316513
step: 30, loss: 0.39431822299957275
step: 40, loss: 0.18117395043373108
step: 50, loss: 0.30597615242004395
step: 60, loss: 0.4668983221054077
step: 70, loss: 0.2392432987689972
step: 80, loss: 0.38073715567588806
step: 90, loss: 0.24190248548984528
step: 100, loss: 0.3097086250782013
step: 110, loss: 0.3836938142776489
step: 120, loss: 0.3753073811531067

```

```

step: 130, loss: 0.31308218836784363
step: 140, loss: 0.39394092559814453
step: 150, loss: 0.322375625371933
step: 160, loss: 0.18986409902572632
step: 170, loss: 0.23666596412658691
step: 180, loss: 0.3149888217449188
step: 190, loss: 0.31045785546302795
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24322813749313354
step: 10, loss: 0.09974983334541321
step: 20, loss: 0.31452658772468567
step: 30, loss: 0.18467269837856293
step: 40, loss: 0.3095807433128357
step: 50, loss: 0.38066262006759644
step: 60, loss: 0.3764537572860718
step: 70, loss: 0.1567390263080597
step: 80, loss: 0.3963170051574707
step: 90, loss: 0.6984395980834961
step: 100, loss: 0.38442859053611755
step: 110, loss: 0.379847913980484
step: 120, loss: 0.6011233329772949
step: 130, loss: 0.3108121454715729
step: 140, loss: 0.31267836689949036
step: 150, loss: 0.2415102869272232
step: 160, loss: 0.231702521443367
step: 170, loss: 0.45721906423568726
step: 180, loss: 0.31200891733169556
step: 190, loss: 0.24100138247013092
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23824848234653473
step: 10, loss: 0.17153941094875336
step: 20, loss: 0.3138156533241272
step: 30, loss: 0.3768276870250702
step: 40, loss: 0.3189157545566559

```

```

step: 50, loss: 0.09863201528787613
step: 60, loss: 0.3169238567352295
step: 70, loss: 0.3131702244281769
step: 80, loss: 0.38922205567359924
step: 90, loss: 0.16181407868862152
step: 100, loss: 0.23377345502376556
step: 110, loss: 0.23965656757354736
step: 120, loss: 0.3858497440814972
step: 130, loss: 0.5091201066970825
step: 140, loss: 0.38230934739112854
step: 150, loss: 0.3125227093696594
step: 160, loss: 0.2392396181821823
step: 170, loss: 0.38128936290740967
step: 180, loss: 0.24116277694702148
step: 190, loss: 0.1606934517621994
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5238012075424194
step: 10, loss: 0.0939902812242508
step: 20, loss: 0.23999819159507751
step: 30, loss: 0.23816046118736267
step: 40, loss: 0.08641743659973145
step: 50, loss: 0.3169732391834259
step: 60, loss: 0.2353401631116867
step: 70, loss: 0.6257898807525635
step: 80, loss: 0.3858390748500824
step: 90, loss: 0.30852729082107544
step: 100, loss: 0.1684272438287735
step: 110, loss: 0.2462766319513321
step: 120, loss: 0.31299254298210144
step: 130, loss: 0.59456467628479
step: 140, loss: 0.45707598328590393
step: 150, loss: 0.31183093786239624
step: 160, loss: 0.31683599948883057
step: 170, loss: 0.44573265314102173
step: 180, loss: 0.3179309070110321
step: 190, loss: 0.10792335122823715
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3755749464035034
step: 10, loss: 0.3132837414741516
step: 20, loss: 0.5364058017730713
step: 30, loss: 0.3914128541946411
step: 40, loss: 0.3098883032798767
step: 50, loss: 0.5267508029937744
step: 60, loss: 0.4482085108757019
step: 70, loss: 0.38680392503738403
step: 80, loss: 0.38108500838279724
step: 90, loss: 0.4462089240550995
step: 100, loss: 0.3074285387992859
step: 110, loss: 0.44888532161712646
step: 120, loss: 0.16760870814323425
step: 130, loss: 0.24169789254665375
step: 140, loss: 0.37847182154655457
step: 150, loss: 0.38171786069869995
step: 160, loss: 0.24272520840168
step: 170, loss: 0.45019957423210144
step: 180, loss: 0.3127448260784149
step: 190, loss: 0.16537520289421082
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3138693571090698
step: 10, loss: 0.31340309977531433
step: 20, loss: 0.3904689848423004
step: 30, loss: 0.3828601837158203
step: 40, loss: 0.24391895532608032
step: 50, loss: 0.2463771253824234
step: 60, loss: 0.31433191895484924
step: 70, loss: 0.16570688784122467
step: 80, loss: 0.3119893968105316
step: 90, loss: 0.3119727373123169
step: 100, loss: 0.16561035811901093
step: 110, loss: 0.6073722243309021
step: 120, loss: 0.1666337549686432
step: 130, loss: 0.24160778522491455
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.31191200017929077
step: 150, loss: 0.39309370517730713

```

```

step: 160, loss: 0.45422613620758057
step: 170, loss: 0.10362377017736435
step: 180, loss: 0.243395134806633
step: 190, loss: 0.31574851274490356
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3094726800918579
step: 10, loss: 0.245820090174675
step: 20, loss: 0.38227900862693787
step: 30, loss: 0.10119929164648056
step: 40, loss: 0.31357336044311523
step: 50, loss: 0.3763827383518219
step: 60, loss: 0.30647170543670654
step: 70, loss: 0.23687660694122314
step: 80, loss: 0.3827097713947296
step: 90, loss: 0.1635976880788803
step: 100, loss: 0.3069594204425812
step: 110, loss: 0.45362457633018494
step: 120, loss: 0.3146549165248871
step: 130, loss: 0.3120695650577545
step: 140, loss: 0.1719307005405426
step: 150, loss: 0.45509073138237
step: 160, loss: 0.38000789284706116
step: 170, loss: 0.24267658591270447
step: 180, loss: 0.30596911907196045
step: 190, loss: 0.5293044447898865
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.524813711643219
step: 10, loss: 0.23457959294319153
step: 20, loss: 0.24445749819278717
step: 30, loss: 0.3008512556552887
step: 40, loss: 0.6581712961196899
step: 50, loss: 0.30971288681030273
step: 60, loss: 0.17415180802345276
step: 70, loss: 0.24022972583770752

```

```

step: 80, loss: 0.2512974143028259
step: 90, loss: 0.3106546700000763
step: 100, loss: 0.38227319717407227
step: 110, loss: 0.4579765498638153
step: 120, loss: 0.23979133367538452
step: 130, loss: 0.3085542321205139
step: 140, loss: 0.457892507314682
step: 150, loss: 0.1669955551624298
step: 160, loss: 0.30821433663368225
step: 170, loss: 0.7421625852584839
step: 180, loss: 0.31096020340919495
step: 190, loss: 0.4618673324584961
/usr/local/lib/python3.7/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=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
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 [00:13, 156.01it/s]

```

```
load_f1 = 0.17216770740410348
real_f1 = 0.17216770740410348
733it [00:00, 3478.43it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:29, 147.33it/s]
```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

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

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

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

Downloading: 100% 481/481 [00:00<00:00, 428kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 741kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 505kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 68.0MB/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: 0.4672192335128784

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

step: 10, loss: 0.48729807138442993

step: 20, loss: 0.34103262424468994

step: 30, loss: 0.4261256754398346

step: 40, loss: 0.5682811141014099

step: 50, loss: 0.29723280668258667

step: 60, loss: 0.5405346155166626

step: 70, loss: 0.31172701716423035

```

step: 80, loss: 0.24045565724372864
step: 90, loss: 0.26209908723831177
step: 100, loss: 0.15724992752075195
step: 110, loss: 0.39969027042388916
step: 120, loss: 0.3266763985157013
step: 130, loss: 0.3260590732097626
step: 140, loss: 0.3815097510814667
step: 150, loss: 0.31938567757606506
step: 160, loss: 0.41484251618385315
step: 170, loss: 0.3169926702976227
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20425090720580616, f1=0.20802534318901794,
best_f1=0.20802534318901794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30589908361434937
step: 10, loss: 0.4448404908180237
step: 20, loss: 0.3112156391143799
step: 30, loss: 0.31806954741477966
step: 40, loss: 0.0791747197508812
step: 50, loss: 0.4159630835056305
step: 60, loss: 0.15798640251159668
step: 70, loss: 0.45975688099861145
step: 80, loss: 0.20900467038154602
step: 90, loss: 0.23683364689350128
step: 100, loss: 0.5130030512809753
step: 110, loss: 0.2812499701976776
step: 120, loss: 0.19722728431224823
step: 130, loss: 0.3040071427822113
step: 140, loss: 0.3296240270137787
step: 150, loss: 0.16980819404125214
step: 160, loss: 0.25304773449897766
step: 170, loss: 0.056788887828588486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6380368098159509, f1=0.69126213592233,

```

```

best_f1=0.69126213592233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4138270914554596
step: 10, loss: 0.09646674990653992
step: 20, loss: 0.08728040009737015
step: 30, loss: 0.09731034189462662
step: 40, loss: 0.11269409954547882
step: 50, loss: 0.42697128653526306
step: 60, loss: 0.1055489331483841
step: 70, loss: 0.14581060409545898
step: 80, loss: 0.08862429857254028
step: 90, loss: 0.4179663062095642
step: 100, loss: 0.10358244925737381
step: 110, loss: 0.13503190875053406
step: 120, loss: 0.07487013190984726
step: 130, loss: 0.21378695964813232
step: 140, loss: 0.10516565293073654
step: 150, loss: 0.17200608551502228
step: 160, loss: 0.14266172051429749
step: 170, loss: 0.06790275126695633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7677261613691931, f1=0.7811764705882352,
best_f1=0.7811764705882352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10575482249259949
step: 10, loss: 0.14450125396251678
step: 20, loss: 0.09532950818538666
step: 30, loss: 0.30634745955467224
step: 40, loss: 0.14696402847766876
step: 50, loss: 0.09427810460329056
step: 60, loss: 0.18037176132202148
step: 70, loss: 0.014072692953050137
step: 80, loss: 0.19442588090896606
step: 90, loss: 0.2706460654735565
step: 100, loss: 0.18295139074325562
step: 110, loss: 0.26449498534202576
step: 120, loss: 0.28323057293891907
step: 130, loss: 0.10363022238016129

```

```

step: 140, loss: 0.11509470641613007
step: 150, loss: 0.14340205490589142
step: 160, loss: 0.010265958495438099
step: 170, loss: 0.052399877458810806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7864077669902912, f1=0.8275862068965516,
best_f1=0.8275862068965516
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009345692582428455
step: 10, loss: 0.04929455369710922
step: 20, loss: 0.025145897641777992
step: 30, loss: 0.17215408384799957
step: 40, loss: 0.15065710246562958
step: 50, loss: 0.08257075399160385
step: 60, loss: 0.045915085822343826
step: 70, loss: 0.01134319044649601
step: 80, loss: 0.013391150161623955
step: 90, loss: 0.07387571781873703
step: 100, loss: 0.03206634148955345
step: 110, loss: 0.15856468677520752
step: 120, loss: 0.1157931387424469
step: 130, loss: 0.016341907903552055
step: 140, loss: 0.07090359926223755
step: 150, loss: 0.07600545138120651
step: 160, loss: 0.10388289391994476
step: 170, loss: 0.07550442218780518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8062015503875968, f1=0.8321513002364066,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19923202693462372
step: 10, loss: 0.06167226657271385

```

```

step: 20, loss: 0.14541618525981903
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.006348264869302511
step: 40, loss: 0.042429011315107346
step: 50, loss: 0.008865343406796455
step: 60, loss: 0.039297137409448624
step: 70, loss: 0.006816979963332415
step: 80, loss: 0.015560387633740902
step: 90, loss: 0.13160039484500885
step: 100, loss: 0.018229825422167778
step: 110, loss: 0.3595142960548401
step: 120, loss: 0.040792860090732574
step: 130, loss: 0.1999596208333969
step: 140, loss: 0.1793462485074997
step: 150, loss: 0.057082679122686386
step: 160, loss: 0.11272105574607849
step: 170, loss: 0.09308585524559021
/usr/local/lib/python3.7/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=0.7901234567901234, f1=0.819672131147541,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06192537769675255
step: 10, loss: 0.1420585662126541
step: 20, loss: 0.01777740754187107
step: 30, loss: 0.0020520808175206184
step: 40, loss: 0.000594179262407124
step: 50, loss: 0.0023286270443350077
step: 60, loss: 0.023235704749822617
step: 70, loss: 0.025172485038638115
step: 80, loss: 0.11646761745214462
step: 90, loss: 0.07945968210697174
step: 100, loss: 0.00906707439571619
step: 110, loss: 0.07664425671100616
step: 120, loss: 0.0252921674400568
step: 130, loss: 0.20542603731155396
step: 140, loss: 0.07481446862220764
step: 150, loss: 0.1264495849609375
step: 160, loss: 0.021074706688523293
step: 170, loss: 0.04687286168336868
/usr/local/lib/python3.7/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=0.7814910025706941, f1=0.8599508599508598,
best_f1=0.8321513002364066

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039940040558576584
step: 10, loss: 0.2719052731990814
step: 20, loss: 0.01761995628476143
step: 30, loss: 0.002252291887998581
step: 40, loss: 0.015055985189974308
step: 50, loss: 0.005038065370172262
step: 60, loss: 0.01156932394951582
step: 70, loss: 0.02366100624203682
step: 80, loss: 0.09172394871711731
step: 90, loss: 0.08160076290369034
step: 100, loss: 0.0038790209218859673
step: 110, loss: 0.09404436498880386
step: 120, loss: 0.1376902163028717
step: 130, loss: 0.010086506605148315
step: 140, loss: 0.03120785392820835
step: 150, loss: 0.2306985855102539
step: 160, loss: 0.01797492988407612
step: 170, loss: 0.09978629648685455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.782608695652174, f1=0.8264840182648402,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04717211425304413
step: 10, loss: 0.010963850654661655
step: 20, loss: 0.0403197780251503
step: 30, loss: 0.005282239522784948
step: 40, loss: 0.18464834988117218
step: 50, loss: 0.002671623369678855
step: 60, loss: 0.11618278175592422
step: 70, loss: 0.014258991926908493
step: 80, loss: 0.004658410791307688
step: 90, loss: 0.023793121799826622
step: 100, loss: 0.0033689760603010654
step: 110, loss: 0.011650036089122295
step: 120, loss: 0.03019869700074196
step: 130, loss: 0.002563304267823696
step: 140, loss: 0.025685252621769905
step: 150, loss: 0.1160295382142067
step: 160, loss: 0.005253904964774847
step: 170, loss: 0.0038734725676476955
/usr/local/lib/python3.7/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=0.7849462365591398, f1=0.8688946015424165,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016092639416456223
step: 10, loss: 0.03135291859507561
step: 20, loss: 0.01190523337572813
step: 30, loss: 0.043323222547769547
step: 40, loss: 0.003052804619073868
step: 50, loss: 0.004207992926239967
step: 60, loss: 0.051224298775196075
step: 70, loss: 0.049743764102458954
step: 80, loss: 0.008980807848274708
step: 90, loss: 0.002165416954085231
step: 100, loss: 0.02932014875113964
step: 110, loss: 0.008940700441598892
step: 120, loss: 0.0019872765988111496
step: 130, loss: 0.000864122761413455
step: 140, loss: 0.0059248595498502254
step: 150, loss: 0.008961142040789127
step: 160, loss: 0.008413732051849365
step: 170, loss: 0.07082128524780273
/usr/local/lib/python3.7/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=0.798941798941799, f1=0.8528678304239402,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09627137333154678
step: 10, loss: 0.003109975252300501
step: 20, loss: 0.0032509882003068924
step: 30, loss: 0.007596049457788467
step: 40, loss: 0.0034168101847171783
step: 50, loss: 0.06407768279314041
step: 60, loss: 0.016621610149741173
step: 70, loss: 0.011823533102869987
step: 80, loss: 0.02202775701880455
step: 90, loss: 0.006636394653469324
step: 100, loss: 0.021819351240992546
step: 110, loss: 0.003515009069815278
step: 120, loss: 0.010005099698901176
step: 130, loss: 0.0026523678097873926
step: 140, loss: 0.005818583536893129

```

```

step: 150, loss: 0.0006761254626326263
step: 160, loss: 0.0188401248306036
step: 170, loss: 0.0688961073756218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8073878627968338, f1=0.8693467336683417,
best_f1=0.8693467336683417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016184797510504723
step: 10, loss: 0.017652079463005066
step: 20, loss: 0.001615923480130732
step: 30, loss: 0.04252452403306961
step: 40, loss: 0.001468677306547761
step: 50, loss: 0.0039320397190749645
step: 60, loss: 0.07837516069412231
step: 70, loss: 0.0007424241630360484
step: 80, loss: 0.0007186861475929618
step: 90, loss: 0.02871713601052761
step: 100, loss: 0.012738167308270931
step: 110, loss: 0.09523127228021622
step: 120, loss: 0.09741254895925522
step: 130, loss: 0.005344010423868895
step: 140, loss: 0.00298051699064672
step: 150, loss: 0.14328975975513458
step: 160, loss: 0.09559953212738037
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.17060783505439758
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8064516129032259, f1=0.8585858585858587,
best_f1=0.8693467336683417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006042541470378637
step: 10, loss: 0.0005155642284080386
step: 20, loss: 0.0029040281660854816
step: 30, loss: 0.10175110399723053
step: 40, loss: 0.002343638800084591
step: 50, loss: 0.009550216607749462

```



```

step: 60, loss: 0.0033940004650503397
step: 70, loss: 0.08649802953004837
step: 80, loss: 0.0017472563777118921
step: 90, loss: 0.004106948617845774
step: 100, loss: 0.046169836074113846
step: 110, loss: 0.004619707819074392
step: 120, loss: 0.0035542533732950687
step: 130, loss: 0.0008713947026990354
step: 140, loss: 0.053696293383836746
step: 150, loss: 0.030225032940506935
step: 160, loss: 0.004133911803364754
step: 170, loss: 0.005567527841776609
/usr/local/lib/python3.7/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=0.8049382716049382, f1=0.8457943925233645,
best_f1=0.8693467336683417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012005766853690147
step: 10, loss: 0.010901996865868568
step: 20, loss: 0.00047496292972937226
step: 30, loss: 0.007165668532252312
step: 40, loss: 0.0008465286227874458
step: 50, loss: 0.0009089545928873122
step: 60, loss: 0.0013752811355516315
step: 70, loss: 0.0008291722042486072
step: 80, loss: 0.0003194195742253214
step: 90, loss: 0.00021988309163134545
step: 100, loss: 0.001100313849747181
step: 110, loss: 0.015036283060908318
step: 120, loss: 0.003781402250751853
step: 130, loss: 0.01408968772739172
step: 140, loss: 0.016316229477524757
step: 150, loss: 0.01788146421313286
step: 160, loss: 0.004474498797208071
step: 170, loss: 0.053568284958601
/usr/local/lib/python3.7/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=0.8009708737864077, f1=0.8411214953271028,
best_f1=0.8693467336683417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005805774708278477
step: 10, loss: 0.00242467038333416

```

```

step: 20, loss: 0.0025935545563697815
step: 30, loss: 0.001184627995826304
step: 40, loss: 0.0009327299194410443
step: 50, loss: 0.0494588166475296
step: 60, loss: 0.0009083157638087869
step: 70, loss: 0.02787487767636776
step: 80, loss: 0.000942667422350496
step: 90, loss: 0.001295579131692648
step: 100, loss: 0.029763080179691315
step: 110, loss: 0.00032400921918451786
step: 120, loss: 0.0024274468887597322
step: 130, loss: 0.0024355968926101923
step: 140, loss: 0.012171491980552673
step: 150, loss: 0.00043206100235693157
step: 160, loss: 0.0006940790917724371
step: 170, loss: 0.034312326461076736
/usr/local/lib/python3.7/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=0.8, f1=0.8550368550368549, best_f1=0.8693467336683417
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:08, 219.25it/s]
load_f1 = 0.45690834473324216
real_f1 = 0.4180645161290322
0it [00:00, ?it/s]/usr/local/lib/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 [00:33, 132.67it/s]

```

4.3 DITTO DIRTY

4.3.1 DBLP-ACM - Running the matcher

```

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

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

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

```

```
--checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.5607209205627441
```

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

```
step: 10, loss: 0.4509860575199127
```

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

```
step: 20, loss: 0.529110312461853
```

```
step: 30, loss: 0.2258455455303192
```

```
step: 40, loss: 0.3706377446651459
```

```
step: 50, loss: 0.6596828699111938
```

```
step: 60, loss: 0.3757329285144806
```

```
step: 70, loss: 0.15582723915576935
```

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

```
step: 80, loss: 0.22769692540168762
```

```
step: 90, loss: 0.15535962581634521
```

```
step: 100, loss: 0.1924682855606079
```

```
step: 110, loss: 0.2543676197528839
```

```
step: 120, loss: 0.10912124067544937
```

```
step: 130, loss: 0.07011356949806213
```

```
step: 140, loss: 0.039748139679431915
```

```
step: 150, loss: 0.266466349363327
```

```

step: 160, loss: 0.08516067266464233
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.2309577763080597
step: 180, loss: 0.1301438957452774
step: 190, loss: 0.10783741623163223
step: 200, loss: 0.029542215168476105
step: 210, loss: 0.1408422440290451
step: 220, loss: 0.11844874173402786
step: 230, loss: 0.007206002715975046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9632107023411371, f1=0.9561304836895389,
best_f1=0.9561304836895389
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024528739973902702
step: 10, loss: 0.2698327898979187
step: 20, loss: 0.06780214607715607
step: 30, loss: 0.03368081524968147
step: 40, loss: 0.09818635135889053
step: 50, loss: 0.024305500090122223
step: 60, loss: 0.011887955479323864
step: 70, loss: 0.017282310873270035
step: 80, loss: 0.0035618736874312162
step: 90, loss: 0.07212669402360916
step: 100, loss: 0.014086905866861343
step: 110, loss: 0.03812975436449051
step: 120, loss: 0.07185718417167664
step: 130, loss: 0.023239240050315857
step: 140, loss: 0.004849319811910391
step: 150, loss: 0.06429637968540192
step: 160, loss: 0.03373449668288231
step: 170, loss: 0.005888691172003746
step: 180, loss: 0.012782194651663303
step: 190, loss: 0.008122287690639496
step: 200, loss: 0.021293900907039642
step: 210, loss: 0.005102457012981176
step: 220, loss: 0.019499430432915688
step: 230, loss: 0.002409757813438773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9655937846836848, f1=0.9655937846836848,
best_f1=0.9655937846836848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022431572899222374
step: 10, loss: 0.005081449635326862
step: 20, loss: 0.06273617595434189
step: 30, loss: 0.006347379181534052
step: 40, loss: 0.05417170375585556
step: 50, loss: 0.022876061499118805
step: 60, loss: 0.02338152565062046
step: 70, loss: 0.0029024165123701096
step: 80, loss: 0.08542889356613159
step: 90, loss: 0.007375055458396673
step: 100, loss: 0.001999608241021633
step: 110, loss: 0.03802531212568283
step: 120, loss: 0.0021527281496673822
step: 130, loss: 0.028441742062568665
step: 140, loss: 0.010540925897657871
step: 150, loss: 0.04889436438679695
step: 160, loss: 0.009491083212196827
step: 170, loss: 0.003304568585008383
step: 180, loss: 0.01990601420402527
step: 190, loss: 0.08006329834461212
step: 200, loss: 0.011402443051338196
step: 210, loss: 0.0030158276204019785
step: 220, loss: 0.022719841450452805
step: 230, loss: 0.0077377017587423325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753363228699552, f1=0.9740698985343857,
best_f1=0.9740698985343857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029262849129736423
step: 10, loss: 0.0011436433997005224
step: 20, loss: 0.007332255132496357

```

```

step: 30, loss: 0.000985784106887877
step: 40, loss: 0.03214186802506447
step: 50, loss: 0.008596604689955711
step: 60, loss: 0.1844097077846527
step: 70, loss: 0.16415104269981384
step: 80, loss: 0.14996236562728882
step: 90, loss: 0.06633000820875168
step: 100, loss: 0.09005885571241379
step: 110, loss: 0.012432116083800793
step: 120, loss: 0.0210232213139534
step: 130, loss: 0.033897515386343
step: 140, loss: 0.005239156074821949
step: 150, loss: 0.010797381401062012
step: 160, loss: 0.014129888266324997
step: 170, loss: 0.001160160987637937
step: 180, loss: 0.15289916098117828
step: 190, loss: 0.005830139387398958
step: 200, loss: 0.04814800247550011
step: 210, loss: 0.008357064798474312
step: 220, loss: 0.006571716163307428
step: 230, loss: 0.0014312237035483122
/usr/local/lib/python3.7/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=0.9668874172185431, f1=0.9709821428571428,
best_f1=0.9740698985343857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019423641497269273
step: 10, loss: 0.03515559062361717
step: 20, loss: 0.13496074080467224
step: 30, loss: 0.003679657354950905
step: 40, loss: 0.009019014425575733
step: 50, loss: 0.0014593605883419514
step: 60, loss: 0.02874971367418766
step: 70, loss: 0.0016585784032940865
step: 80, loss: 0.033380404114723206
step: 90, loss: 0.06701623648405075
step: 100, loss: 0.00044331650133244693
step: 110, loss: 0.007971706800162792
step: 120, loss: 0.001277224742807448
step: 130, loss: 0.00038915409822948277
step: 140, loss: 0.007823776453733444
step: 150, loss: 0.024650927633047104
step: 160, loss: 0.0014911438338458538
step: 170, loss: 0.004896119236946106
step: 180, loss: 0.003877254668623209

```

```

step: 190, loss: 0.10848107188940048
step: 200, loss: 0.027851207181811333
step: 210, loss: 0.011324985884130001
step: 220, loss: 0.0016611971659585834
step: 230, loss: 0.002512772800400853
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9762174405436014,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001390518737025559
step: 10, loss: 0.001200521015562117
step: 20, loss: 0.001099185785278678
step: 30, loss: 0.023395827040076256
step: 40, loss: 0.0004459224292077124
step: 50, loss: 0.003059315262362361
step: 60, loss: 0.0010206984588876367
step: 70, loss: 0.14984571933746338
step: 80, loss: 0.0032814722508192062
step: 90, loss: 0.008199404925107956
step: 100, loss: 0.000490249483846128
step: 110, loss: 0.06014486774802208
step: 120, loss: 0.0010955638717859983
step: 130, loss: 0.0011285132495686412
step: 140, loss: 0.0005160594009794295
step: 150, loss: 0.00033329991856589913
step: 160, loss: 0.003561606863513589
step: 170, loss: 0.007776949089020491
step: 180, loss: 0.013112717308104038
step: 190, loss: 0.003114887047559023
step: 200, loss: 0.020977918058633804
step: 210, loss: 0.00272384169511497
step: 220, loss: 0.0016514656599611044
step: 230, loss: 0.0007371939718723297
/usr/local/lib/python3.7/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=0.9711751662971175, f1=0.96875, best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.0027371873147785664
step: 10, loss: 0.0006889420910738409
step: 20, loss: 0.0006824673037044704
step: 30, loss: 0.0003075319982599467
step: 40, loss: 0.0009708765428513288
step: 50, loss: 0.0013742039445787668
step: 60, loss: 0.001112740719690919
step: 70, loss: 0.0004745943006128073
step: 80, loss: 0.00030245634843595326
step: 90, loss: 0.00642688712105155
step: 100, loss: 0.0005110340425744653
step: 110, loss: 0.0004059879283886403
step: 120, loss: 0.0007349016377702355
step: 130, loss: 0.0017004951369017363
step: 140, loss: 0.00016152142779901624
step: 150, loss: 0.03499981760978699
step: 160, loss: 0.00025494524743407965
step: 170, loss: 0.00033584923949092627
step: 180, loss: 0.00023494775814469904
step: 190, loss: 0.006164866499602795
step: 200, loss: 0.0734722763299942
step: 210, loss: 0.004322807304561138
step: 220, loss: 0.0022731549106538296
step: 230, loss: 0.003076343797147274
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.976271186440678,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005014349590055645
step: 10, loss: 0.03926306590437889
step: 20, loss: 0.0008470373577438295
step: 30, loss: 0.0013696140376850963
step: 40, loss: 0.001044840202666819
step: 50, loss: 0.0023770735133439302
step: 60, loss: 0.0007677511312067509
step: 70, loss: 0.0010941133368760347
step: 80, loss: 0.02845616079866886
step: 90, loss: 0.0015716180205345154
step: 100, loss: 0.0004916632315143943
step: 110, loss: 0.0020418092608451843
step: 120, loss: 0.0003098730812780559
step: 130, loss: 0.01870287023484707
step: 140, loss: 0.00014006967830937356
step: 150, loss: 0.061403319239616394

```

```

step: 160, loss: 0.002726609818637371
step: 170, loss: 0.0005655785789713264
step: 180, loss: 0.0008515493827871978
step: 190, loss: 0.08020959794521332
step: 200, loss: 0.00296444701962173
step: 210, loss: 0.0021803565323352814
step: 220, loss: 0.0004460023483261466
step: 230, loss: 0.0003089826786890626
/usr/local/lib/python3.7/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=0.976324689966178, f1=0.9749430523917996,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011312509886920452
step: 10, loss: 0.00011439032095950097
step: 20, loss: 0.0006648043636232615
step: 30, loss: 0.0002215206332039088
step: 40, loss: 0.00015684918616898358
step: 50, loss: 0.00015459705900866538
step: 60, loss: 0.0008190696244128048
step: 70, loss: 0.04297546669840813
step: 80, loss: 0.0002935132069978863
step: 90, loss: 0.007928687147796154
step: 100, loss: 0.00014646310592070222
step: 110, loss: 9.248020069207996e-05
step: 120, loss: 0.02540801279246807
step: 130, loss: 0.000722517492249608
step: 140, loss: 0.005880812648683786
step: 150, loss: 0.00014704419299960136
step: 160, loss: 0.0006958949961699545
step: 170, loss: 5.909970059292391e-05
step: 180, loss: 0.0018002393189817667
step: 190, loss: 6.354922516038641e-05
step: 200, loss: 0.02603604644536972
step: 210, loss: 0.009039044380187988
step: 220, loss: 9.687652345746756e-05
step: 230, loss: 0.15883159637451172
/usr/local/lib/python3.7/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=0.9680264608599779, f1=0.9700332963374029,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00011847222049254924
step: 10, loss: 0.00040273796184919775
step: 20, loss: 0.0002689194807317108
step: 30, loss: 0.0012266701087355614
step: 40, loss: 0.0005953277577646077
step: 50, loss: 0.00012078943836968392
step: 60, loss: 0.0003854628885164857
step: 70, loss: 0.003382593160495162
step: 80, loss: 0.009141565300524235
step: 90, loss: 8.149415953084826e-05
step: 100, loss: 0.001370133482851088
step: 110, loss: 0.002084042178466916
step: 120, loss: 0.0002689658722374588
step: 130, loss: 0.0011235696729272604
step: 140, loss: 0.00012819314724765718
step: 150, loss: 0.0013852391857653856
step: 160, loss: 3.315393041702919e-05
step: 170, loss: 0.0002982545120175928
step: 180, loss: 0.007078028749674559
step: 190, loss: 0.0005252551054582
step: 200, loss: 0.00024846597807481885
step: 210, loss: 0.0012334099737927318
step: 220, loss: 0.00030151064856909215
step: 230, loss: 0.0010448388056829572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9776286353467561, f1=0.9750566893424036,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007046031532809138
step: 10, loss: 0.000556106329895556
step: 20, loss: 0.0006460556178353727
step: 30, loss: 0.00012483549653552473
step: 40, loss: 4.491562503972091e-05
step: 50, loss: 0.0013947271509096026
step: 60, loss: 0.0006157017778605223
step: 70, loss: 6.207117985468358e-05
step: 80, loss: 0.012112668715417385
step: 90, loss: 0.024548353627324104
step: 100, loss: 0.0007834061980247498
step: 110, loss: 0.00018489800277166069
step: 120, loss: 0.0016803968464955688
step: 130, loss: 0.00010656366794137284
step: 140, loss: 0.0002943750296253711
step: 150, loss: 0.00025467705563642085

```

```

step: 160, loss: 0.0318891741335392
step: 170, loss: 0.0012813342036679387
step: 180, loss: 0.0018755351193249226
step: 190, loss: 0.00023564582807011902
step: 200, loss: 0.014368844218552113
step: 210, loss: 0.00019848474767059088
step: 220, loss: 0.05277812108397484
step: 230, loss: 0.000200386974029243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.978675645342312, f1=0.9694915254237287,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002572540193796158
step: 10, loss: 0.00012712737952824682
step: 20, loss: 0.003972481470555067
step: 30, loss: 0.007911760360002518
step: 40, loss: 0.00031098019098863006
step: 50, loss: 0.0059194788336753845
step: 60, loss: 0.0003558439202606678
step: 70, loss: 4.535133484750986e-05
step: 80, loss: 2.835155464708805e-05
step: 90, loss: 0.0018175089498981833
step: 100, loss: 5.439554661279544e-05
step: 110, loss: 3.969618774135597e-05
step: 120, loss: 0.07449933141469955
step: 130, loss: 0.01178064476698637
step: 140, loss: 0.00010363920591771603
step: 150, loss: 0.0028264313004910946
step: 160, loss: 0.0010706533212214708
step: 170, loss: 8.678841550135985e-05
step: 180, loss: 0.017123915255069733
step: 190, loss: 0.0042048548348248005
step: 200, loss: 4.2029274482047185e-05
step: 210, loss: 0.02098572812974453
step: 220, loss: 0.0019279983825981617
step: 230, loss: 0.00062794319819659
/usr/local/lib/python3.7/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=0.9787234042553192, f1=0.9808342728297633,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0004095916519872844
step: 10, loss: 0.000586576119530946
step: 20, loss: 4.838140739593655e-05
step: 30, loss: 0.00011228855873923749
step: 40, loss: 0.0003072126128245145
step: 50, loss: 0.0014400780200958252
step: 60, loss: 3.997659223387018e-05
step: 70, loss: 0.0013538802741095424
step: 80, loss: 0.0016039916081354022
step: 90, loss: 3.7520578189287335e-05
step: 100, loss: 8.66452173795551e-05
step: 110, loss: 5.391438025981188e-05
step: 120, loss: 5.069729013484903e-05
step: 130, loss: 6.599365588044748e-05
step: 140, loss: 0.00011090839689131826
step: 150, loss: 3.333908171043731e-05
step: 160, loss: 0.0015568180242553353
step: 170, loss: 3.4877612051786855e-05
step: 180, loss: 0.013966269791126251
step: 190, loss: 0.00016005028737708926
step: 200, loss: 0.00025980910868383944
step: 210, loss: 0.00707454327493906
step: 220, loss: 0.00012583239004015923
step: 230, loss: 0.00013780321751255542
/usr/local/lib/python3.7/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=0.9788182831661093, f1=0.9775280898876404,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.1931402190821245e-05
step: 10, loss: 5.651538958773017e-05
step: 20, loss: 0.0038798709865659475
step: 30, loss: 4.222744610160589e-05
step: 40, loss: 9.233137097908184e-05
step: 50, loss: 2.7099125873064622e-05
step: 60, loss: 4.2610256059560925e-05
step: 70, loss: 8.285455987788737e-05
step: 80, loss: 3.968016608268954e-05
step: 90, loss: 3.933271364076063e-05
step: 100, loss: 0.00034563784720376134
step: 110, loss: 3.100817775703035e-05
step: 120, loss: 2.4425335141131654e-05
step: 130, loss: 0.00014727313828188926
step: 140, loss: 4.804567652172409e-05
step: 150, loss: 5.476999649545178e-05

```

```

step: 160, loss: 3.5650944482767954e-05
step: 170, loss: 5.736559978686273e-05
step: 180, loss: 4.396500298753381e-05
step: 190, loss: 7.80412956373766e-05
step: 200, loss: 9.121008042711765e-05
step: 210, loss: 2.3062337277224287e-05
step: 220, loss: 0.00032411576830781996
step: 230, loss: 0.08894933015108109
/usr/local/lib/python3.7/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=0.9776785714285714, f1=0.978675645342312,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008449605666100979
step: 10, loss: 2.7540098017198034e-05
step: 20, loss: 0.00014134972298052162
step: 30, loss: 8.359939965885133e-05
step: 40, loss: 2.1304278561729006e-05
step: 50, loss: 0.00011138732224935666
step: 60, loss: 0.03850388154387474
step: 70, loss: 0.002049972303211689
step: 80, loss: 0.00015937455464154482
step: 90, loss: 0.00022575966431759298
step: 100, loss: 1.7702162949717604e-05
step: 110, loss: 0.00012338589294813573
step: 120, loss: 0.007382952142506838
step: 130, loss: 0.00017424806719645858
step: 140, loss: 0.0016028059180825949
step: 150, loss: 0.000575756945181638
step: 160, loss: 0.00026146919117309153
step: 170, loss: 7.244003791129217e-05
step: 180, loss: 4.432124114828184e-05
step: 190, loss: 0.0011187541531398892
step: 200, loss: 0.0017706912476569414
step: 210, loss: 0.005620005540549755
step: 220, loss: 2.2965710741118528e-05
step: 230, loss: 0.004598632454872131
/usr/local/lib/python3.7/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=0.9776785714285714, f1=0.9763779527559054,
best_f1=0.9762174405436014
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:17, 140.50it/s]
load_f1 = 0.9808342728297633
real_f1 = 0.9785794813979707
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:33, 133.03it/s]

```

4.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

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

step: 0, loss: 0.641252875328064

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

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

step: 10, loss: 0.46240201592445374

step: 20, loss: 0.262587308883667

step: 30, loss: 0.4109558165073395

step: 40, loss: 0.3871670961380005

step: 50, loss: 0.6471396684646606

step: 60, loss: 0.3284994065761566

step: 70, loss: 0.44418424367904663

step: 80, loss: 0.5266823768615723

step: 90, loss: 0.388607919216156

step: 100, loss: 0.5632707476615906

step: 110, loss: 0.44722914695739746

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

step: 120, loss: 0.37602317333221436

step: 130, loss: 0.3169570863246918

step: 140, loss: 0.22716166079044342

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

step: 150, loss: 0.39893338084220886

step: 160, loss: 0.3016855716705322

step: 170, loss: 0.33969855308532715

step: 180, loss: 0.21071767807006836

step: 190, loss: 0.2516101002693176

step: 200, loss: 0.09253160655498505

step: 210, loss: 0.04548788443207741

step: 220, loss: 0.17408154904842377

step: 230, loss: 0.06581620126962662

step: 240, loss: 0.015137654729187489

step: 250, loss: 0.11629397422075272

step: 260, loss: 0.2651892900466919

step: 270, loss: 0.3906925916671753

step: 280, loss: 0.11369942128658295

step: 290, loss: 0.09269116073846817

step: 300, loss: 0.22664949297904968

step: 310, loss: 0.1193731278181076

step: 320, loss: 0.06436026096343994

step: 330, loss: 0.06968390941619873

step: 340, loss: 0.23137341439723969

step: 350, loss: 0.21875087916851044

step: 360, loss: 0.02703467756509781

step: 370, loss: 0.05479070544242859

step: 380, loss: 0.2029087096452713

step: 390, loss: 0.028268471360206604

```

step: 400, loss: 0.1487249732017517
step: 410, loss: 0.2367410510778427
step: 420, loss: 0.020631346851587296
step: 430, loss: 0.03442563861608505
step: 440, loss: 0.04017849266529083
step: 450, loss: 0.14074969291687012
step: 460, loss: 0.11407289654016495
step: 470, loss: 0.04919647425413132
step: 480, loss: 0.113076351583004
step: 490, loss: 0.25335389375686646
step: 500, loss: 0.08335985243320465
step: 510, loss: 0.11532045900821686
step: 520, loss: 0.28381040692329407
step: 530, loss: 0.03795235604047775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9044352994970278, f1=0.906021897810219,
best_f1=0.906021897810219
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13963249325752258
step: 10, loss: 0.1066543310880661
step: 20, loss: 0.23514465987682343
step: 30, loss: 0.0967787653207779
step: 40, loss: 0.0505545474588871
step: 50, loss: 0.1471806913614273
step: 60, loss: 0.10355575382709503
step: 70, loss: 0.06735482811927795
step: 80, loss: 0.2045045644044876
step: 90, loss: 0.04775141924619675
step: 100, loss: 0.2136431783437729
step: 110, loss: 0.03803636506199837
step: 120, loss: 0.1692710816860199
step: 130, loss: 0.03984066843986511
step: 140, loss: 0.09352327883243561
step: 150, loss: 0.02716651000082493
step: 160, loss: 0.0640101358294487
step: 170, loss: 0.27548977732658386
step: 180, loss: 0.09570490568876266
step: 190, loss: 0.10354001075029373
step: 200, loss: 0.14967875182628632
step: 210, loss: 0.04102381318807602

```

```

step: 220, loss: 0.007918120361864567
step: 230, loss: 0.09910725057125092
step: 240, loss: 0.043552931398153305
step: 250, loss: 0.08265195041894913
step: 260, loss: 0.09429077804088593
step: 270, loss: 0.04242652654647827
step: 280, loss: 0.056556493043899536
step: 290, loss: 0.04886501282453537
step: 300, loss: 0.09021003544330597
step: 310, loss: 0.1691015511751175
step: 320, loss: 0.08577694743871689
step: 330, loss: 0.1385514736175537
step: 340, loss: 0.1866416335105896
step: 350, loss: 0.006719588302075863
step: 360, loss: 0.09614860266447067
step: 370, loss: 0.10054890066385269
step: 380, loss: 0.13504508137702942
step: 390, loss: 0.015253128483891487
step: 400, loss: 0.09658831357955933
step: 410, loss: 0.015099641866981983
step: 420, loss: 0.06080363690853119
step: 430, loss: 0.06897738575935364
step: 440, loss: 0.008929572999477386
step: 450, loss: 0.21576082706451416
step: 460, loss: 0.08277571201324463
step: 470, loss: 0.08788184821605682
step: 480, loss: 0.008168667554855347
step: 490, loss: 0.13878768682479858
step: 500, loss: 0.03948076069355011
step: 510, loss: 0.016170311719179153
step: 520, loss: 0.3238777220249176
step: 530, loss: 0.06698572635650635
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9208371246587808, f1=0.9223826714801444,
best_f1=0.9223826714801444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10376933962106705
step: 10, loss: 0.21353694796562195
step: 20, loss: 0.018762530758976936
step: 30, loss: 0.10172540694475174

```

step: 40, loss: 0.1577981412410736
step: 50, loss: 0.05727578327059746
step: 60, loss: 0.04364883899688721
step: 70, loss: 0.01966862753033638
step: 80, loss: 0.13406378030776978
step: 90, loss: 0.028305796906352043
step: 100, loss: 0.032920386642217636
step: 110, loss: 0.046204689890146255
step: 120, loss: 0.19962239265441895
step: 130, loss: 0.04765962436795235
step: 140, loss: 0.07634016871452332
step: 150, loss: 0.05689239129424095
step: 160, loss: 0.028192149475216866
step: 170, loss: 0.006057855673134327
step: 180, loss: 0.03790688514709473
step: 190, loss: 0.04946554824709892
step: 200, loss: 0.02771974727511406
step: 210, loss: 0.028727874159812927
step: 220, loss: 0.10138525068759918
step: 230, loss: 0.027932047843933105
step: 240, loss: 0.1401033103466034
step: 250, loss: 0.14589931070804596
step: 260, loss: 0.07772228866815567
step: 270, loss: 0.016600124537944794
step: 280, loss: 0.030229290947318077
step: 290, loss: 0.03065469115972519
step: 300, loss: 0.05471101775765419
step: 310, loss: 0.07421493530273438
step: 320, loss: 0.02128017134964466
step: 330, loss: 0.00680353119969368
step: 340, loss: 0.04026905074715614
step: 350, loss: 0.11683405935764313
step: 360, loss: 0.03980594873428345
step: 370, loss: 0.21658530831336975
step: 380, loss: 0.06358473002910614
step: 390, loss: 0.048092011362314224
step: 400, loss: 0.29155561327934265
step: 410, loss: 0.033668503165245056
step: 420, loss: 0.02389342710375786
step: 430, loss: 0.047461118549108505
step: 440, loss: 0.18771792948246002
step: 450, loss: 0.11262819916009903
step: 460, loss: 0.07588577270507812
step: 470, loss: 0.0764971673488617
step: 480, loss: 0.25618940591812134
step: 490, loss: 0.025839442387223244
step: 500, loss: 0.017668871209025383
step: 510, loss: 0.0685914158821106

```

step: 520, loss: 0.009903103113174438
step: 530, loss: 0.007690760772675276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9265750828991001, f1=0.9243856332703214,
best_f1=0.9243856332703214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08043995499610901
step: 10, loss: 0.02323758229613304
step: 20, loss: 0.06087363138794899
step: 30, loss: 0.15954937040805817
step: 40, loss: 0.03897140547633171
step: 50, loss: 0.029160641133785248
step: 60, loss: 0.038906779140233994
step: 70, loss: 0.19720415771007538
step: 80, loss: 0.133699432015419
step: 90, loss: 0.03308963403105736
step: 100, loss: 0.026959998533129692
step: 110, loss: 0.1334039568901062
step: 120, loss: 0.028584718704223633
step: 130, loss: 0.08276471495628357
step: 140, loss: 0.07630795240402222
step: 150, loss: 0.034755606204271317
step: 160, loss: 0.02567538432776928
step: 170, loss: 0.016913093626499176
step: 180, loss: 0.048586778342723846
step: 190, loss: 0.032621610909700394
step: 200, loss: 0.042497020214796066
step: 210, loss: 0.04590988904237747
step: 220, loss: 0.056506626307964325
step: 230, loss: 0.022762125357985497
step: 240, loss: 0.001962372101843357
step: 250, loss: 0.0792258083820343
step: 260, loss: 0.004957257770001888
step: 270, loss: 0.25935178995132446
step: 280, loss: 0.010624291375279427
step: 290, loss: 0.004802980925887823
step: 300, loss: 0.00417490815743804
step: 310, loss: 0.026706580072641373
step: 320, loss: 0.11151094734668732
step: 330, loss: 0.07502622157335281

```

```

step: 340, loss: 0.015367239713668823
step: 350, loss: 0.11205384880304337
step: 360, loss: 0.025990601629018784
step: 370, loss: 0.004447086248546839
step: 380, loss: 0.019360434263944626
step: 390, loss: 0.0023945297580212355
step: 400, loss: 0.02375722862780094
step: 410, loss: 0.0070162005722522736
step: 420, loss: 0.0229920856654644
step: 430, loss: 0.03158063068985939
step: 440, loss: 0.007332616485655308
step: 450, loss: 0.045955583453178406
step: 460, loss: 0.01404153648763895
step: 470, loss: 0.005293187219649553
step: 480, loss: 0.010298643261194229
step: 490, loss: 0.0014409363502636552
step: 500, loss: 0.0778718814253807
step: 510, loss: 0.03019845113158226
step: 520, loss: 0.012911185622215271
step: 530, loss: 0.21222969889640808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9343410486537552, f1=0.9325153374233129,
best_f1=0.9325153374233129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014086456038057804
step: 10, loss: 0.030281970277428627
step: 20, loss: 0.01427631638944149
step: 30, loss: 0.08197043091058731
step: 40, loss: 0.004718948621302843
step: 50, loss: 0.1457977592945099
step: 60, loss: 0.05859748646616936
step: 70, loss: 0.0021778088994324207
step: 80, loss: 0.007484584115445614
step: 90, loss: 0.035249803215265274
step: 100, loss: 0.1726476550102234
step: 110, loss: 0.003245042636990547
step: 120, loss: 0.11044630408287048
step: 130, loss: 0.02582353726029396
step: 140, loss: 0.0759657546877861
step: 150, loss: 0.01301409862935543

```

```
step: 160, loss: 0.04986615106463432
step: 170, loss: 0.03891431540250778
step: 180, loss: 0.01201004721224308
step: 190, loss: 0.0066870227456092834
step: 200, loss: 0.011576410382986069
step: 210, loss: 0.003757760627195239
step: 220, loss: 0.008168061263859272
step: 230, loss: 0.007460588589310646
step: 240, loss: 0.006100642029196024
step: 250, loss: 0.14674338698387146
step: 260, loss: 0.003427287796512246
step: 270, loss: 0.006210556719452143
step: 280, loss: 0.10437419265508652
step: 290, loss: 0.0249948650598526
step: 300, loss: 0.292941153049469
step: 310, loss: 0.010875776410102844
step: 320, loss: 0.15631835162639618
step: 330, loss: 0.016208849847316742
step: 340, loss: 0.02780204266309738
step: 350, loss: 0.01814953237771988
step: 360, loss: 0.000899137114174664
step: 370, loss: 0.001427723909728229
step: 380, loss: 0.0013053487055003643
step: 390, loss: 0.044897958636283875
step: 400, loss: 0.007436858024448156
step: 410, loss: 0.15088944137096405
step: 420, loss: 0.17501895129680634
step: 430, loss: 0.022695694118738174
step: 440, loss: 0.003818945959210396
step: 450, loss: 0.006048201117664576
step: 460, loss: 0.06588118523359299
step: 470, loss: 0.038363017141819
step: 480, loss: 0.012735245749354362
step: 490, loss: 0.02310566045343876
step: 500, loss: 0.04052351042628288
step: 510, loss: 0.01763072796165943
step: 520, loss: 0.0355418436229229
step: 530, loss: 0.05192413926124573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9357798165137614, f1=0.9304467987102718,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

step: 0, loss: 0.02572234719991684

step: 10, loss: 0.007143519353121519

step: 20, loss: 0.0051306369714438915

step: 30, loss: 0.0009735539788380265

step: 40, loss: 0.009511376731097698

step: 50, loss: 0.0010856487788259983

step: 60, loss: 0.01798117160797119

step: 70, loss: 0.004439658019691706

step: 80, loss: 0.00536199938505888

step: 90, loss: 0.0064674122259020805

step: 100, loss: 0.023115677759051323

step: 110, loss: 0.024328477680683136

step: 120, loss: 0.015577500686049461

step: 130, loss: 0.0008253763080574572

step: 140, loss: 0.0008717464515939355

step: 150, loss: 0.0006403761799447238

step: 160, loss: 0.02782703936100006

step: 170, loss: 0.0012882621958851814

step: 180, loss: 0.027320513501763344

step: 190, loss: 0.09484854340553284

step: 200, loss: 0.018504761159420013

step: 210, loss: 0.007549296133220196

step: 220, loss: 0.024420883506536484

step: 230, loss: 0.001805932610295713

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

step: 240, loss: 0.005904574412852526

step: 250, loss: 0.08521228283643723

step: 260, loss: 0.01954953372478485

step: 270, loss: 0.001783955143764615

step: 280, loss: 0.061384137719869614

step: 290, loss: 0.006842795293778181

step: 300, loss: 0.007261286955326796

step: 310, loss: 0.13623712956905365

step: 320, loss: 0.0011944080470129848

step: 330, loss: 0.01049035880714655

step: 340, loss: 0.003842149395495653

step: 350, loss: 0.005366834346204996

step: 360, loss: 0.056079987436532974

step: 370, loss: 0.0363074354827404

step: 380, loss: 0.0034371677320450544

step: 390, loss: 0.007718069478869438

step: 400, loss: 0.002479083137586713

step: 410, loss: 0.0005608664359897375

step: 420, loss: 0.0018635377055034041

step: 430, loss: 0.010569781996309757

step: 440, loss: 0.025878995656967163


```
step: 450, loss: 0.13964290916919708
step: 460, loss: 0.0015400080010294914
step: 470, loss: 0.02038388140499592
step: 480, loss: 0.009416443295776844
step: 490, loss: 0.006610884331166744
step: 500, loss: 0.0021887626498937607
step: 510, loss: 0.15665942430496216
step: 520, loss: 0.011668269522488117
step: 530, loss: 0.023180175572633743
/usr/local/lib/python3.7/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=0.9327058823529412, f1=0.925035360678925,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006488660350441933
step: 10, loss: 0.0020253786351531744
step: 20, loss: 0.08781320601701736
step: 30, loss: 0.01007561944425106
step: 40, loss: 0.010619660839438438
step: 50, loss: 0.034421395510435104
step: 60, loss: 0.013754408806562424
step: 70, loss: 0.01589159667491913
step: 80, loss: 0.004658241756260395
step: 90, loss: 0.009617878124117851
step: 100, loss: 0.0012700125807896256
step: 110, loss: 0.0023688459768891335
step: 120, loss: 0.14064738154411316
step: 130, loss: 0.05749762803316116
step: 140, loss: 0.006787050981074572
step: 150, loss: 0.00963220652192831
step: 160, loss: 0.027052177116274834
step: 170, loss: 0.001681141322478652
step: 180, loss: 0.036390457302331924
step: 190, loss: 0.044064074754714966
step: 200, loss: 0.0004116872441954911
step: 210, loss: 0.0021834997460246086
step: 220, loss: 0.023907089605927467
step: 230, loss: 0.009419772773981094
step: 240, loss: 0.010847224853932858
step: 250, loss: 0.1632813960313797
step: 260, loss: 0.0015395191730931401
step: 270, loss: 0.003962475806474686
step: 280, loss: 0.06248900666832924
step: 290, loss: 0.026717908680438995
step: 300, loss: 0.0009759168024174869
```

```

step: 310, loss: 0.0021342216059565544
step: 320, loss: 0.022164365276694298
step: 330, loss: 0.005458793602883816
step: 340, loss: 0.010744141414761543
step: 350, loss: 0.0012907690834254026
step: 360, loss: 0.09670531749725342
step: 370, loss: 0.000806672964245081
step: 380, loss: 0.0073914737440645695
step: 390, loss: 0.0025678297970443964
step: 400, loss: 0.07279433310031891
step: 410, loss: 0.0009156010928563774
step: 420, loss: 0.04235388711094856
step: 430, loss: 0.003202216699719429
step: 440, loss: 0.0037299655377864838
step: 450, loss: 0.004562329966574907
step: 460, loss: 0.028954284265637398
step: 470, loss: 0.19219210743904114
step: 480, loss: 0.004639994818717241
step: 490, loss: 0.009494648315012455
step: 500, loss: 0.004906816408038139
step: 510, loss: 0.0023354538716375828
step: 520, loss: 0.16404619812965393
step: 530, loss: 0.011908830143511295
/usr/local/lib/python3.7/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=0.9336405529953917, f1=0.930939226519337,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006285510607995093
step: 10, loss: 0.0032299067825078964
step: 20, loss: 0.003019599709659815
step: 30, loss: 0.015460584312677383
step: 40, loss: 0.0017206204356625676
step: 50, loss: 0.01844790391623974
step: 60, loss: 0.002473424421623349
step: 70, loss: 0.0016025465447455645
step: 80, loss: 0.003678024746477604
step: 90, loss: 0.0028526834212243557
step: 100, loss: 0.025837639346718788
step: 110, loss: 0.00010540212679188699
step: 120, loss: 0.00032268051290884614
step: 130, loss: 0.003643299452960491
step: 140, loss: 0.09754901379346848
step: 150, loss: 0.005058083683252335
step: 160, loss: 0.00180948490742594

```

```

step: 170, loss: 0.010104479268193245
step: 180, loss: 0.005488782189786434
step: 190, loss: 0.14042183756828308
step: 200, loss: 0.026159094646573067
step: 210, loss: 0.015667513012886047
step: 220, loss: 0.021492721512913704
step: 230, loss: 0.129616841673851
step: 240, loss: 0.02789035066962242
step: 250, loss: 0.0037321033887565136
step: 260, loss: 0.0015100479358807206
step: 270, loss: 0.0012488627107813954
step: 280, loss: 0.000410392734920606
step: 290, loss: 0.019820550456643105
step: 300, loss: 0.005505036097019911
step: 310, loss: 0.03027394600212574
step: 320, loss: 0.0027936610858887434
step: 330, loss: 0.0022526662796735764
step: 340, loss: 0.02332121692597866
step: 350, loss: 0.001609751139767468
step: 360, loss: 0.017291205003857613
step: 370, loss: 0.11268136650323868
step: 380, loss: 0.00016775184485595673
step: 390, loss: 0.0039099655114114285
step: 400, loss: 0.0037823666352778673
step: 410, loss: 0.11067462712526321
step: 420, loss: 0.000684052356518805
step: 430, loss: 0.02227775566279888
step: 440, loss: 0.08927096426486969
step: 450, loss: 0.0028223709668964148
step: 460, loss: 0.04453081265091896
step: 470, loss: 0.12014646828174591
step: 480, loss: 0.032692696899175644
step: 490, loss: 0.0047144717536866665
step: 500, loss: 0.0006209408165886998
step: 510, loss: 0.016415879130363464
step: 520, loss: 0.0013047511456534266
step: 530, loss: 0.0029641343280673027
/usr/local/lib/python3.7/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=0.9344413665743306, f1=0.9312413474850022,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009620062191970646
step: 10, loss: 0.009480256587266922
step: 20, loss: 0.0031124302186071873

```

step: 30, loss: 0.02270159311592579
step: 40, loss: 0.000830767210572958
step: 50, loss: 0.0008722839411348104
step: 60, loss: 0.0019530735444277525
step: 70, loss: 0.005331294145435095
step: 80, loss: 0.016957467421889305
step: 90, loss: 0.0576326809823513
step: 100, loss: 0.007061779033392668
step: 110, loss: 0.004418172407895327
step: 120, loss: 0.014752894639968872
step: 130, loss: 0.0021184347569942474
step: 140, loss: 0.0013968657003715634
step: 150, loss: 0.02265273779630661
step: 160, loss: 0.005562334321439266
step: 170, loss: 0.002983545884490013
step: 180, loss: 0.008712958544492722
step: 190, loss: 0.0029185451567173004
step: 200, loss: 0.0038299274165183306
step: 210, loss: 0.00026220784639008343
step: 220, loss: 0.001474374788813293
step: 230, loss: 0.0005225192289799452
step: 240, loss: 0.0010639207903295755
step: 250, loss: 0.002632708754390478
step: 260, loss: 0.0011191020021215081
step: 270, loss: 0.021499712020158768
step: 280, loss: 0.005297649651765823
step: 290, loss: 0.0007904420490376651
step: 300, loss: 0.028278181329369545
step: 310, loss: 0.14823265373706818
step: 320, loss: 0.0006587272509932518
step: 330, loss: 0.0008977138786576688
step: 340, loss: 0.017220929265022278
step: 350, loss: 0.06855326890945435
step: 360, loss: 0.015517204999923706
step: 370, loss: 0.16570134460926056
step: 380, loss: 0.0020090187899768353
step: 390, loss: 0.0007311197696253657
step: 400, loss: 0.043710820376873016
step: 410, loss: 0.00842322502285242
step: 420, loss: 0.0002218126319348812
step: 430, loss: 0.0038559348322451115
step: 440, loss: 0.00010938812920358032
step: 450, loss: 0.005449840798974037
step: 460, loss: 0.006944888737052679
step: 470, loss: 0.001281590899452567
step: 480, loss: 9.593598952051252e-05
step: 490, loss: 0.0030905005987733603
step: 500, loss: 0.06615836173295975

```

step: 510, loss: 0.0017089609755203128
step: 520, loss: 0.004008363466709852
step: 530, loss: 0.0035003814846277237
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9351981351981352, f1=0.9314685314685315,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010185926221311092
step: 10, loss: 0.0033448601607233286
step: 20, loss: 0.053409475833177567
step: 30, loss: 0.0007277077529579401
step: 40, loss: 0.00038455953472293913
step: 50, loss: 0.0003008010098710656
step: 60, loss: 0.0021110938396304846
step: 70, loss: 0.010142212733626366
step: 80, loss: 0.04454091191291809
step: 90, loss: 0.0012480991426855326
step: 100, loss: 0.0020566764287650585
step: 110, loss: 0.12432076036930084
step: 120, loss: 0.0736037865281105
step: 130, loss: 0.017032671719789505
step: 140, loss: 0.0030713253654539585
step: 150, loss: 0.012333403341472149
step: 160, loss: 0.0027323749382048845
step: 170, loss: 0.0007506864494644105
step: 180, loss: 0.014550134539604187
step: 190, loss: 0.05263494327664375
step: 200, loss: 0.0001481577637605369
step: 210, loss: 0.0021281521767377853
step: 220, loss: 8.332246216014028e-05
step: 230, loss: 0.0027549031656235456
step: 240, loss: 0.019572528079152107
step: 250, loss: 0.009111425839364529
step: 260, loss: 0.0041082254610955715
step: 270, loss: 0.060314614325761795
step: 280, loss: 0.050652991980314255
step: 290, loss: 0.003758950624614954
step: 300, loss: 0.03137316554784775
step: 310, loss: 0.008922236040234566
step: 320, loss: 0.0426403246819973
step: 330, loss: 0.004328755661845207
step: 340, loss: 0.0009373885113745928
step: 350, loss: 0.010054416954517365
step: 360, loss: 0.0018079594010487199

```

```

step: 370, loss: 0.0015759046655148268
step: 380, loss: 0.0005197604768909514
step: 390, loss: 0.0005724822985939682
step: 400, loss: 0.001127026043832302
step: 410, loss: 0.0018588397651910782
step: 420, loss: 0.00013704333105124533
step: 430, loss: 0.022366447374224663
step: 440, loss: 0.00018806164734996855
step: 450, loss: 0.07319329679012299
step: 460, loss: 0.0007616705843247473
step: 470, loss: 0.014940241351723671
step: 480, loss: 0.06609850376844406
step: 490, loss: 0.02870851755142212
step: 500, loss: 0.0015700969379395247
step: 510, loss: 0.0015494631370529532
step: 520, loss: 0.0002481910341884941
step: 530, loss: 0.004419103730469942
/usr/local/lib/python3.7/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=0.9357277882797731, f1=0.9262062440870389,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009640366770327091
step: 10, loss: 0.002105213236063719
step: 20, loss: 0.0003368695033714175
step: 30, loss: 0.0071396604180336
step: 40, loss: 0.005012327805161476
step: 50, loss: 0.0001573883491801098
step: 60, loss: 0.005544999614357948
step: 70, loss: 7.717406697338447e-05
step: 80, loss: 7.899715274106711e-05
step: 90, loss: 0.0001998517254833132
step: 100, loss: 0.0016816123388707638
step: 110, loss: 1.2379054169286974e-05
step: 120, loss: 0.008165468461811543
step: 130, loss: 3.644844764494337e-05
step: 140, loss: 0.021963413804769516
step: 150, loss: 0.0021398006938397884
step: 160, loss: 0.08378429710865021
step: 170, loss: 0.005982758477330208
step: 180, loss: 0.0005837592761963606
step: 190, loss: 0.013978201895952225
step: 200, loss: 0.001867663930170238
step: 210, loss: 0.002069758018478751
step: 220, loss: 0.05283267796039581

```

```

step: 230, loss: 0.0010473651345819235
step: 240, loss: 0.00034943711943924427
step: 250, loss: 0.00022982984955888242
step: 260, loss: 0.006967087741941214
step: 270, loss: 0.000872421485837549
step: 280, loss: 0.0016646296717226505
step: 290, loss: 0.012850035913288593
step: 300, loss: 1.5783789422130212e-05
step: 310, loss: 0.000352597504388541
step: 320, loss: 0.000865796348080039
step: 330, loss: 2.6604722734191455e-05
step: 340, loss: 0.1013365313410759
step: 350, loss: 0.0008992115617729723
step: 360, loss: 0.0014495272189378738
step: 370, loss: 0.0005525229498744011
step: 380, loss: 0.0017837488558143377
step: 390, loss: 0.0005496472585946321
step: 400, loss: 9.908715583151206e-05
step: 410, loss: 0.00024246842076536268
step: 420, loss: 0.00010380869935033843
step: 430, loss: 0.002004479756578803
step: 440, loss: 0.008585313335061073
step: 450, loss: 0.002014677505940199
step: 460, loss: 0.00428543658927083
step: 470, loss: 0.0015383537393063307
step: 480, loss: 0.001680539920926094
step: 490, loss: 0.009850016795098782
step: 500, loss: 0.0033226090017706156
step: 510, loss: 0.0035815718583762646
step: 520, loss: 0.012678670696914196
step: 530, loss: 0.0011335789458826184
/usr/local/lib/python3.7/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=0.9327770050996754, f1=0.9304267161410019,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004752339154947549
step: 10, loss: 0.0009221700602211058
step: 20, loss: 0.003857053117826581
step: 30, loss: 0.002569349482655525
step: 40, loss: 0.00014931611076463014
step: 50, loss: 0.0008423435501754284
step: 60, loss: 0.0005141167785041034
step: 70, loss: 0.0023323444183915854
step: 80, loss: 0.001253334223292768

```

```
step: 90, loss: 0.00536379124969244
step: 100, loss: 0.32793471217155457
step: 110, loss: 0.10288690030574799
step: 120, loss: 0.0012599111068993807
step: 130, loss: 0.0021408130414783955
step: 140, loss: 0.00040104350773617625
step: 150, loss: 0.000780395872425288
step: 160, loss: 0.003567769890651107
step: 170, loss: 0.0006145237712189555
step: 180, loss: 0.0006361962878145278
step: 190, loss: 0.00019137997878715396
step: 200, loss: 0.002756700851023197
step: 210, loss: 0.001564333913847804
step: 220, loss: 0.0013714480446651578
step: 230, loss: 0.0003358772664796561
step: 240, loss: 0.001089761033654213
step: 250, loss: 3.091543112532236e-05
step: 260, loss: 0.0012025312753394246
step: 270, loss: 0.04240390658378601
step: 280, loss: 0.001203348976559937
step: 290, loss: 0.02003549039363861
step: 300, loss: 0.01093356590718031
step: 310, loss: 0.0003754920035135001
step: 320, loss: 0.0037479118909686804
step: 330, loss: 0.0037179547362029552
step: 340, loss: 0.000819999782834202
step: 350, loss: 0.001587655395269394
step: 360, loss: 0.004885923117399216
step: 370, loss: 0.0034540831111371517
step: 380, loss: 3.578432369977236e-05
step: 390, loss: 0.011368805542588234
step: 400, loss: 0.00017660539015196264
step: 410, loss: 0.0016497650649398565
step: 420, loss: 0.13096053898334503
step: 430, loss: 0.001371844788081944
step: 440, loss: 0.0006610113778151572
step: 450, loss: 0.0018359494861215353
step: 460, loss: 0.0023329302202910185
step: 470, loss: 0.01640579104423523
step: 480, loss: 0.002526394557207823
step: 490, loss: 0.002207942772656679
step: 500, loss: 0.002951685106381774
step: 510, loss: 0.006452379748225212
step: 520, loss: 0.012055093422532082
step: 530, loss: 0.002940306905657053
/usr/local/lib/python3.7/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=0.9348729792147806, f1=0.9299539170506912,
best_f1=0.9304467987102718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
 warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003021145472303033
step: 10, loss: 0.005511550232768059
step: 20, loss: 0.0014474405907094479
step: 30, loss: 0.007721361704170704
step: 40, loss: 0.021201424300670624
step: 50, loss: 0.005221855361014605
step: 60, loss: 0.0015349297318607569
step: 70, loss: 0.0039132763631641865
step: 80, loss: 0.0016231924528256059
step: 90, loss: 0.0002780678041744977
step: 100, loss: 0.0018284769030287862
step: 110, loss: 0.0003964229836128652
step: 120, loss: 0.0007008523098193109
step: 130, loss: 0.0003906900528818369
step: 140, loss: 0.06514175981283188
step: 150, loss: 0.00042040395783260465
step: 160, loss: 0.017254674807190895
step: 170, loss: 0.0007127219578251243
step: 180, loss: 0.0017720831092447042
step: 190, loss: 0.03433346003293991
step: 200, loss: 0.0001846089435275644
step: 210, loss: 0.00017388757260050625
step: 220, loss: 0.0003641262010205537
step: 230, loss: 0.001187002519145608
step: 240, loss: 0.00033230058033950627
step: 250, loss: 0.00022905903460923582
step: 260, loss: 0.004479446914047003
step: 270, loss: 0.03587828204035759
step: 280, loss: 0.0009065764024853706
step: 290, loss: 0.0005752643919549882
step: 300, loss: 0.0005439961096271873
step: 310, loss: 0.0019068318651989102
step: 320, loss: 0.004095497541129589
step: 330, loss: 0.0009213233133777976
step: 340, loss: 0.007308322004973888
step: 350, loss: 0.000475710432510823
step: 360, loss: 0.09795411676168442
step: 370, loss: 0.001259076059795916
step: 380, loss: 0.0015813540667295456
step: 390, loss: 0.00040611677104607224
step: 400, loss: 0.004868134390562773
step: 410, loss: 0.0005561565048992634
step: 420, loss: 0.00047958645154722035

```

step: 430, loss: 0.000631383212748915
step: 440, loss: 0.060906726866960526
step: 450, loss: 0.06377125531435013
step: 460, loss: 0.004550032317638397
step: 470, loss: 0.004415615927428007
step: 480, loss: 0.00020786962704733014
step: 490, loss: 0.005681054666638374
step: 500, loss: 0.0069024693220853806
step: 510, loss: 0.0002398017095401883
step: 520, loss: 0.001293362583965063
step: 530, loss: 0.00027288240380585194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9365079365079366, f1=0.9338919925512105,
best_f1=0.9338919925512105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017180710565298796
step: 10, loss: 0.004092342220246792
step: 20, loss: 0.0049938359297811985
step: 30, loss: 0.0007514338940382004
step: 40, loss: 0.00033587831421755254
step: 50, loss: 0.00037222960963845253
step: 60, loss: 0.00021806874428875744
step: 70, loss: 0.008133492432534695
step: 80, loss: 0.0003960456815548241
step: 90, loss: 5.3032475989311934e-05
step: 100, loss: 0.017690757289528847
step: 110, loss: 0.000493004044983536
step: 120, loss: 3.466036287136376e-05
step: 130, loss: 0.00014752749120816588
step: 140, loss: 0.011035200208425522
step: 150, loss: 0.08905836939811707
step: 160, loss: 0.003103996394202113
step: 170, loss: 0.001953926868736744
step: 180, loss: 0.0038384387735277414
step: 190, loss: 0.0044641452841460705
step: 200, loss: 0.0044436706230044365
step: 210, loss: 0.01011980976909399
step: 220, loss: 4.44254583271686e-05
step: 230, loss: 0.00010604610724840313
step: 240, loss: 0.00010293714149156585

```

```

step: 250, loss: 0.0028795262333005667
step: 260, loss: 0.0015592498239129782
step: 270, loss: 0.020078051835298538
step: 280, loss: 0.00025368380011059344
step: 290, loss: 0.00025662360712885857
step: 300, loss: 0.00026068894658237696
step: 310, loss: 0.04468139633536339
step: 320, loss: 0.0017615616088733077
step: 330, loss: 0.0008539105183444917
step: 340, loss: 0.0001764461339917034
step: 350, loss: 3.6348683352116495e-05
step: 360, loss: 0.006513491272926331
step: 370, loss: 0.009285630658268929
step: 380, loss: 0.005366365425288677
step: 390, loss: 0.001434289151802659
step: 400, loss: 0.0019508452387526631
step: 410, loss: 0.0001432677818229422
step: 420, loss: 0.0009711617021821439
step: 430, loss: 0.006066967733204365
step: 440, loss: 0.04410339519381523
step: 450, loss: 0.0005780683713965118
step: 460, loss: 0.08413222432136536
step: 470, loss: 1.0389740054961294e-05
step: 480, loss: 0.004143682308495045
step: 490, loss: 0.00037305918522179127
step: 500, loss: 0.0009676121990196407
step: 510, loss: 0.027028165757656097
step: 520, loss: 0.003101535839959979
step: 530, loss: 0.0020185529720038176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9369966840360018, f1=0.9310018903591682,
best_f1=0.9310018903591682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005081783165223897
step: 10, loss: 0.0007970698643475771
step: 20, loss: 0.0018460000865161419
step: 30, loss: 0.001425385009497404
step: 40, loss: 0.00023887820134405047
step: 50, loss: 0.0034821571316570044
step: 60, loss: 0.0037815412506461143

```

step: 70, loss: 8.968626207206398e-05
step: 80, loss: 0.002920864149928093
step: 90, loss: 0.0007926154648885131
step: 100, loss: 0.00010384162305854261
step: 110, loss: 0.0006630782736465335
step: 120, loss: 0.0004174677887931466
step: 130, loss: 0.013155517168343067
step: 140, loss: 0.0001589038292877376
step: 150, loss: 0.002524935407564044
step: 160, loss: 3.665835174615495e-05
step: 170, loss: 6.955364369787276e-05
step: 180, loss: 0.001794784446246922
step: 190, loss: 0.0022194108460098505
step: 200, loss: 0.03948375582695007
step: 210, loss: 0.0078001199290156364
step: 220, loss: 0.10845708847045898
step: 230, loss: 0.013613028451800346
step: 240, loss: 0.000145033496664837
step: 250, loss: 0.000245054776314646
step: 260, loss: 0.0001831705740187317
step: 270, loss: 0.00018510106019675732
step: 280, loss: 2.099705488944892e-05
step: 290, loss: 0.004286462441086769
step: 300, loss: 0.001342741772532463
step: 310, loss: 0.024254946038126945
step: 320, loss: 0.0017000912921503186
step: 330, loss: 4.226374585414305e-05
step: 340, loss: 0.0009988852543756366
step: 350, loss: 0.0033646065276116133
step: 360, loss: 0.007892395369708538
step: 370, loss: 0.0011869373265653849
step: 380, loss: 0.0004552532627712935
step: 390, loss: 0.058538176119327545
step: 400, loss: 0.04722920060157776
step: 410, loss: 0.0026965716388076544
step: 420, loss: 0.0012508039362728596
step: 430, loss: 6.115116411820054e-05
step: 440, loss: 0.004144568927586079
step: 450, loss: 0.003850491251796484
step: 460, loss: 0.0017816723557189107
step: 470, loss: 2.660285827005282e-05
step: 480, loss: 0.0014015864580869675
step: 490, loss: 0.00025032093981280923
step: 500, loss: 0.0030320819932967424
step: 510, loss: 0.014343425631523132
step: 520, loss: 0.0026361129712313414
step: 530, loss: 0.0003273691690992564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9378531073446328, f1=0.9325210871602625,
best_f1=0.9325210871602625
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 [00:35, 160.52it/s]
load_f1 = 0.9330210772833724
real_f1 = 0.9332087809434843
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 [00:31, 138.16it/s]

```