

EMedium_90_1_2_bert

August 13, 2022

0.1 Install packages

```
[ ]: !git clone https://github.com/megagonlabs/ditto
      %cd ditto
      !pip install -r requirements.txt
      !git clone https://github.com/rit-git/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 | 16.96 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 1.3 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 11.5 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 62.8 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.2 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 63.9 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 23.6 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 | 10.19 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 17.6 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
|                                     | 195 kB 74.6 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 7.1 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 3.6 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 48.0 MB/s
Collecting jsonlines==1.2.0
  Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)

```

```

Collecting nltk==3.4.5
  Downloading nltk-3.4.5.zip (1.5 MB)
    | 1.5 MB 57.8 MB/s
Requirement already satisfied: scipy>=0.18.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.7.3)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 65.3 MB/s
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 49.0 MB/s
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: 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)
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 blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
    | 3.7 MB 60.2 MB/s
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
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: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-

```

```

packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    |                                     | 880 kB 53.6 MB/s
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    |                                     | 3.0 MB 53.1 MB/s
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
    Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449924
    sha256=94ed9485e223d8d47ae5a6c5a7a820a963c22fcb9d56ba26ad6b1cdf2260048
    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=cadc26b1a4f4fbabccf7078bce34e21e47c80b1284706c9f812f84e9d9546e21
  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: 10117, done.
remote: Counting objects: 100% (233/233), done.
remote: Compressing objects: 100% (165/165), done.
remote: Total 10117 (delta 119), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.97 MiB | 17.52 MiB/s, done.
Resolving deltas: 100% (6927/6927), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writable
Created temporary directory: /tmp/pip-ephem-wheel-cache-n7v7xz8c
Created temporary directory: /tmp/pip-req-tracker-bom6i7ro
Initialized build tracking at /tmp/pip-req-tracker-bom6i7ro
Created build tracker: /tmp/pip-req-tracker-bom6i7ro
Entered build tracker: /tmp/pip-req-tracker-bom6i7ro
```

```

Created temporary directory: /tmp/pip-install-ru1tn5oc
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-i_nf88da
  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-
bom6i7ro'
  Running setup.py (path:/tmp/pip-req-build-i_nf88da/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-md2gxes9
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-md2gxes9/apex.egg-info
  writing /tmp/pip-pip-egg-info-md2gxes9/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-md2gxes9/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-md2gxes9/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-md2gxes9/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-md2gxes9/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-i_nf88da 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-bom6i7ro'
Created temporary directory: /tmp/pip-unpack-pfl3_qke
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-2g4ey3s_
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-2g4ey3s_
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-i_nf88da/setup.py''';

```



```
__file__='''/tmp/pip-req-build-i_nf88da/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-2g4ey3s_
```

```
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/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.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
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
creating build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
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/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
creating build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adagrad.py -> 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_lamb.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
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/handle.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/opt.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
creating build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
copying apex/normalization/__init__.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/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
creating build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/_timers.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/utils.py ->
build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/__init__.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/_data
  copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
  copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
  creating build/lib/apex/transformer/testing
  copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/__init__.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
  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/tensor_parallel
  copying apex/transformer/tensor_parallel/mappings.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/data.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/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
  creating build/lib/apex/transformer/amp
  copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp

```

```

    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_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
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.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/clip_grad
    copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    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/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->

```

```

build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/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_encdec_multihead_attn_norm_add_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
    creating build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
    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/index_mul_2d
  copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
  copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
  creating build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
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_memory.py ->
build/lib/apex/contrib/peer_memory
  creating build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
  creating build/lib/apex/contrib/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
  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/__init__.py ->
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
  creating build/lib/apex/amp/lists
  copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
  copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
  copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
  copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
  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/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/sync_batchnorm_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
  copying build/lib/apex/parallel/multiproc.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/__init__.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
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    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/transformer
    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/transformer/pipeline_parallel
    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/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_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/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel

```

```

    copying build/lib/apex/transformer/pipeline_parallel/__init__.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/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
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
    creating build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/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/testing/standalone_bert.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/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/microbatches.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
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/memory.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/data.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/cross_entropy.py ->

```



```

build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating 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/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating 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_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/__init__.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
  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
  creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/__init__.py ->
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
  creating build/bdist.linux-x86_64/wheel/apex/contrib
  creating 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/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/test.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/__init__.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_module_test.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
  creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
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
  creating 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
  copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  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/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/fast_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
  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_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn

```

```

    copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating 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
    copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    creating 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/call_permut
ation_search_kernels.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/__init__.py
->
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/sparse_masklib.py ->
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/asp.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
    creating build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.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_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers

```

```

    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
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    creating build/bdist.linux-x86_64/wheel/apex/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/__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_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating 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/__version__.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/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_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/opt.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/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
    copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp

```

```

    copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    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/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
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
    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-2g4ey3s_/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
    adding 'apex/__init__.py'
    adding 'apex/_autocast_utils.py'
    adding 'apex/RNN/RNNBackend.py'

```

```

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

```

```

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

```

```

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

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298244
sha256=8d6b0c47a5a2cd9022c7ea214b37c28c613fc7933fd92e24ef556f80a0c47904
Stored in directory: /tmp/pip-ephem-wheel-cache-n7v7xz8c/wheels/6b/22/a2/16655
26ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-bom6i7ro'
/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 32.1 MB/s
Collecting awscli
  Downloading awscli-1.25.49-py3-none-any.whl (3.9 MB)
    |                                     | 3.9 MB 60.1 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                                     | 79 kB 8.1 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting botocore==1.27.49
  Downloading botocore-1.27.49-py3-none-any.whl (9.0 MB)
    |                                     | 9.0 MB 63.7 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                                     | 548 kB 51.2 MB/s
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.49->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==1.27.49->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-1.25.49 botocore-1.27.49 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 16 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.

torchtex 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: 1054, done.
remote: Counting objects: 100% (124/124), done.
remote: Compressing objects: 100% (91/91), done.
remote: Total 1054 (delta 51), reused 46 (delta 21), pack-reused 930
Receiving objects: 100% (1054/1054), 257.86 MiB | 17.32 MiB/s, done.
Resolving deltas: 100% (635/635), done.
Checking out files: 100% (1304/1304), done.

```

0.4 remove e move

```
[ ]: !ls
```

```

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

```

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

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

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

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

Downloading: 100% 433/433 [00:00<00:00, 361kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 1.22MB/s]

Downloading: 100% 440M/440M [00:09<00:00, 47.3MB/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: 0.5486657023429871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3146067415730337, f1=0.28571428571428575,
best_f1=0.28571428571428575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48944440484046936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5, f1=0.45283018867924535, best_f1=0.45283018867924535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5245822668075562
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.6000000000000001, 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: 0.2815271317958832
/usr/local/lib/python3.7/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=0.6000000000000001, f1=0.43902439024390244, 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: 0.14917969703674316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.4864864864864865,
best_f1=0.4864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2067754715681076
/usr/local/lib/python3.7/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=0.6, f1=0.6666666666666666, best_f1=0.4864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2823704779148102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6956521739130435, f1=0.6923076923076924,
best_f1=0.6923076923076924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024911433458328247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7407407407407408, f1=0.689655172413793,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008418473415076733
/usr/local/lib/python3.7/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=0.7407407407407408, f1=0.6875000000000001,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0590173713862896
/usr/local/lib/python3.7/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=0.7200000000000001, f1=0.6666666666666666,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006879309192299843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7586206896551724, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0195121169090271
/usr/local/lib/python3.7/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=0.7586206896551724, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032519202679395676
/usr/local/lib/python3.7/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=0.75, f1=0.6923076923076924, best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00924671906977892
/usr/local/lib/python3.7/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=0.75, f1=0.6923076923076924, best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016090277582406998
/usr/local/lib/python3.7/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=0.75, f1=0.6923076923076924, best_f1=0.6875000000000001
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, 114241.74it/s]
/usr/local/lib/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 = 0.7692307692307692
real_f1 = 0.7407407407407408
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 250.20it/s]

```

1.1.2 DBLP-ACM - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:

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

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

```
step: 0, loss: 0.6050437092781067
step: 10, loss: 0.5961460471153259
step: 20, loss: 0.3053739666938782
step: 30, loss: 0.17074282467365265
step: 40, loss: 0.1738070249557495
step: 50, loss: 0.030076224356889725
step: 60, loss: 0.13243995606899261
step: 70, loss: 0.1342955231666565
step: 80, loss: 0.21282190084457397
step: 90, loss: 0.16349375247955322
step: 100, loss: 0.009060807526111603
step: 110, loss: 0.07085710018873215
step: 120, loss: 0.017651362344622612
step: 130, loss: 0.007159739267081022
step: 140, loss: 0.002775422530248761
step: 150, loss: 0.04542158171534538
step: 160, loss: 0.0034491107799112797
step: 170, loss: 0.019006308168172836
step: 180, loss: 0.006152716930955648
step: 190, loss: 0.029655691236257553
step: 200, loss: 0.0035077440552413464
step: 210, loss: 0.02644363045692444
step: 220, loss: 0.02246606908738613
step: 230, loss: 0.01890481263399124
```

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

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9645714285714285, f1=0.9643268124280783,
best_f1=0.9643268124280783
```

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

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013306129723787308
step: 10, loss: 0.009337375871837139
step: 20, loss: 0.16054180264472961
step: 30, loss: 0.11397097259759903
step: 40, loss: 0.012058273889124393
step: 50, loss: 0.004575543571263552
```

```

step: 60, loss: 0.004986582323908806
step: 70, loss: 0.20065777003765106
step: 80, loss: 0.003245992586016655
step: 90, loss: 0.018155453726649284
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.013951990753412247
step: 110, loss: 0.008598334155976772
step: 120, loss: 0.0024457161780446768
step: 130, loss: 0.0016631667967885733
step: 140, loss: 0.005418557208031416
step: 150, loss: 0.02326616272330284
step: 160, loss: 0.028149481862783432
step: 170, loss: 0.0012279716320335865
step: 180, loss: 0.0033935671672225
step: 190, loss: 0.005310466978698969
step: 200, loss: 0.001765528111718595
step: 210, loss: 0.0007536948542110622
step: 220, loss: 0.1842622607946396
step: 230, loss: 0.0592818520963192
/usr/local/lib/python3.7/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=0.9380952380952382, f1=0.9342891278375149,
best_f1=0.9643268124280783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.137958362698555
step: 10, loss: 0.11054724454879761
step: 20, loss: 0.006394913420081139
step: 30, loss: 0.05833986774086952
step: 40, loss: 0.10199776291847229
step: 50, loss: 0.0046935477294027805
step: 60, loss: 0.0031134248711168766
step: 70, loss: 0.0027690064162015915
step: 80, loss: 0.0004492285952437669
step: 90, loss: 0.13894328474998474
step: 100, loss: 0.0004428801476024091
step: 110, loss: 0.00580056244507432
step: 120, loss: 0.006741393357515335
step: 130, loss: 0.001300302566960454
step: 140, loss: 0.0017881226958706975
step: 150, loss: 0.0018877164693549275
step: 160, loss: 0.036424268037080765
step: 170, loss: 0.00982165988534689
step: 180, loss: 0.01716361567378044
step: 190, loss: 0.00045654361019842327
step: 200, loss: 0.050820980221033096

```

```

step: 210, loss: 0.0003909154038410634
step: 220, loss: 0.00021857218234799802
step: 230, loss: 0.00020256753487046808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987598647125141, f1=0.9841628959276018,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020726947695948184
step: 10, loss: 0.00018373533384874463
step: 20, loss: 0.00023979424440767616
step: 30, loss: 0.00015096389688551426
step: 40, loss: 0.015958864241838455
step: 50, loss: 0.0006793386419303715
step: 60, loss: 0.0005664818454533815
step: 70, loss: 0.00035367393866181374
step: 80, loss: 0.005092989653348923
step: 90, loss: 0.0003840524877887219
step: 100, loss: 0.0006028374191373587
step: 110, loss: 0.0003600159543566406
step: 120, loss: 0.03417307883501053
step: 130, loss: 0.00030023197177797556
step: 140, loss: 0.00013886578381061554
step: 150, loss: 0.0016143054235726595
step: 160, loss: 8.455825445707887e-05
step: 170, loss: 0.009453731589019299
step: 180, loss: 0.0007103837560862303
step: 190, loss: 0.0022717081010341644
step: 200, loss: 0.0004207464517094195
step: 210, loss: 0.05279038846492767
step: 220, loss: 0.00030604866333305836
step: 230, loss: 0.003308597719296813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876265466816648, f1=0.9808773903262092,
best_f1=0.9808773903262092

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006085021886974573
step: 10, loss: 0.00035447272239252925
step: 20, loss: 0.0005576703115366399
step: 30, loss: 0.0012799593387171626
step: 40, loss: 0.0001841510966187343
step: 50, loss: 0.00015561029431410134
step: 60, loss: 0.0003146347007714212
step: 70, loss: 0.00011874691699631512
step: 80, loss: 0.00015707689453847706
step: 90, loss: 0.00010279324487783015
step: 100, loss: 0.0001307336351601407
step: 110, loss: 9.205925016431138e-05
step: 120, loss: 6.509260856546462e-05
step: 130, loss: 0.00038617514655925333
step: 140, loss: 0.00031430405215360224
step: 150, loss: 0.00015621459169778973
step: 160, loss: 0.0011178338900208473
step: 170, loss: 0.00026640109717845917
step: 180, loss: 0.0012940340675413609
step: 190, loss: 0.00028331318753771484
step: 200, loss: 0.00024945265613496304
step: 210, loss: 0.04733138903975487
step: 220, loss: 0.010957603342831135
step: 230, loss: 0.0005319140036590397
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9737742303306728,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007283852901309729
step: 10, loss: 0.00023453222820535302
step: 20, loss: 0.017088349908590317
step: 30, loss: 0.001000956050120294
step: 40, loss: 0.00041253832750953734
step: 50, loss: 0.011859684251248837
step: 60, loss: 0.00016745425818953663
step: 70, loss: 0.0005525251617655158
step: 80, loss: 0.0002919239632319659
step: 90, loss: 0.000411965767852962
step: 100, loss: 0.07057476788759232
step: 110, loss: 0.0006588955293409526
step: 120, loss: 0.0002596937119960785

```

```

step: 130, loss: 0.0001966497366083786
step: 140, loss: 0.00015000748680904508
step: 150, loss: 0.0001897243782877922
step: 160, loss: 0.003724890761077404
step: 170, loss: 0.0005319913616403937
step: 180, loss: 0.04435988515615463
step: 190, loss: 0.053711287677288055
step: 200, loss: 0.00015291750605683774
step: 210, loss: 0.0003528543747961521
step: 220, loss: 0.0005861306563019753
step: 230, loss: 0.005862909369170666
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9783352337514253,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020066031720489264
step: 10, loss: 0.00022313893714454025
step: 20, loss: 0.0013416935689747334
step: 30, loss: 0.00177520583383739
step: 40, loss: 0.00024187425151467323
step: 50, loss: 0.0002797571651171893
step: 60, loss: 0.0013494286686182022
step: 70, loss: 0.009696432389318943
step: 80, loss: 0.00016399536980316043
step: 90, loss: 8.444456761935726e-05
step: 100, loss: 0.00014064341667108238
step: 110, loss: 0.0010539263021200895
step: 120, loss: 0.0001388130767736584
step: 130, loss: 0.00011032017937395722
step: 140, loss: 0.00022019314928911626
step: 150, loss: 0.00011612293747020885
step: 160, loss: 0.0040192860178649426
step: 170, loss: 0.0003013988898601383
step: 180, loss: 0.0012340714456513524
step: 190, loss: 0.00013954816677141935
step: 200, loss: 0.2060476839542389
step: 210, loss: 0.00015318169607780874
step: 220, loss: 0.0010345714399591088
step: 230, loss: 0.0004746485792566091
/usr/local/lib/python3.7/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=0.9863945578231292, f1=0.975, best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001198364479932934
step: 10, loss: 0.00041992837213911116
step: 20, loss: 0.0001155403588199988
step: 30, loss: 0.00015510470257140696
step: 40, loss: 0.0028890189714729786
step: 50, loss: 0.00032417188049294055
step: 60, loss: 0.0001259749406017363
step: 70, loss: 0.0006449136999435723
step: 80, loss: 0.0019055560696870089
step: 90, loss: 0.0028921670746058226
step: 100, loss: 0.00010098711936734617
step: 110, loss: 0.0032412749715149403
step: 120, loss: 0.0001479465136071667
step: 130, loss: 0.00010362044849898666
step: 140, loss: 5.700138717656955e-05
step: 150, loss: 0.0001483097585150972
step: 160, loss: 8.43002344481647e-05
step: 170, loss: 6.011184086673893e-05
step: 180, loss: 0.00011413671018090099
step: 190, loss: 0.0002882153494283557
step: 200, loss: 5.737453466281295e-05
step: 210, loss: 5.7105859013972804e-05
step: 220, loss: 0.00017470672901254147
step: 230, loss: 7.787172944517806e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9853768278965129, f1=0.9752252252252253,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.655599569668993e-05
step: 10, loss: 0.00014794959861319512
step: 20, loss: 0.0011025918647646904
step: 30, loss: 0.0001286031329073012
step: 40, loss: 0.0016456771409139037
step: 50, loss: 6.374122312990949e-05
step: 60, loss: 0.00032562727574259043
step: 70, loss: 0.0007615704089403152
step: 80, loss: 0.00015574763529002666
step: 90, loss: 0.0003854703390970826
step: 100, loss: 0.0003676922933664173
step: 110, loss: 8.937346137827262e-05
step: 120, loss: 3.916265632142313e-05
step: 130, loss: 0.0002680830948520452

```

```

step: 140, loss: 3.6920006095897406e-05
step: 150, loss: 0.0004932312876917422
step: 160, loss: 0.00010766398918349296
step: 170, loss: 0.00015696971968282014
step: 180, loss: 0.0006885085022076964
step: 190, loss: 0.0011740918271243572
step: 200, loss: 0.00052775809308514
step: 210, loss: 0.0032374097499996424
step: 220, loss: 0.00011055549111915752
step: 230, loss: 0.00025941431522369385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9876543209876544, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028690582257695496
step: 10, loss: 0.0001167413211078383
step: 20, loss: 0.00010459381155669689
step: 30, loss: 6.0423335526138544e-05
step: 40, loss: 4.8485249863006175e-05
step: 50, loss: 6.944559572730213e-05
step: 60, loss: 9.031651279656217e-05
step: 70, loss: 0.0008680394385010004
step: 80, loss: 4.776764762937091e-05
step: 90, loss: 0.0011162408627569675
step: 100, loss: 4.9959824536927044e-05
step: 110, loss: 4.0000515582505614e-05
step: 120, loss: 6.900980952195823e-05
step: 130, loss: 4.1810406401054934e-05
step: 140, loss: 0.009603803977370262
step: 150, loss: 0.00033094428363256156
step: 160, loss: 3.7724174035247415e-05
step: 170, loss: 3.3757314668037e-05
step: 180, loss: 0.0032563921995460987
step: 190, loss: 0.0022240306716412306
step: 200, loss: 4.018490653834306e-05
step: 210, loss: 5.8622499636840075e-05
step: 220, loss: 4.498762064031325e-05
step: 230, loss: 5.48749158042483e-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=0.9898305084745763, f1=0.9841628959276018,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.209998249076307e-05
step: 10, loss: 8.169359352905303e-05
step: 20, loss: 0.0008705875952728093
step: 30, loss: 0.034364230930805206
step: 40, loss: 4.404069477459416e-05
step: 50, loss: 3.438333806116134e-05
step: 60, loss: 4.306187474867329e-05
step: 70, loss: 0.00012119331950088963
step: 80, loss: 3.3008767786668614e-05
step: 90, loss: 3.1670941098127514e-05
step: 100, loss: 3.812888462562114e-05
step: 110, loss: 0.0017974593210965395
step: 120, loss: 4.010077827842906e-05
step: 130, loss: 7.096199260558933e-05
step: 140, loss: 5.966125172562897e-05
step: 150, loss: 0.03322030231356621
step: 160, loss: 2.9782999263261445e-05
step: 170, loss: 0.001389636192470789
step: 180, loss: 0.0001414179860148579
step: 190, loss: 3.5953773476649076e-05
step: 200, loss: 0.0003581395430956036
step: 210, loss: 3.094893691013567e-05
step: 220, loss: 2.9957991500850767e-05
step: 230, loss: 0.00013698582188226283
/usr/local/lib/python3.7/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=0.9898074745186863, f1=0.9818594104308391,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.693167647928931e-05
step: 10, loss: 2.9290691600181162e-05
step: 20, loss: 2.875474456232041e-05
step: 30, loss: 5.202598913456313e-05
step: 40, loss: 3.995202496298589e-05
step: 50, loss: 0.0002691402623895556

```

```

step: 60, loss: 0.00020432825840543956
step: 70, loss: 3.4653661714401096e-05
step: 80, loss: 7.796953286742792e-05
step: 90, loss: 2.7163867343915626e-05
step: 100, loss: 3.980129986302927e-05
step: 110, loss: 4.523159077507444e-05
step: 120, loss: 3.675160769489594e-05
step: 130, loss: 2.5089406335609965e-05
step: 140, loss: 3.468113936833106e-05
step: 150, loss: 4.91248065372929e-05
step: 160, loss: 2.667624175956007e-05
step: 170, loss: 3.497516081552021e-05
step: 180, loss: 0.0004405809741001576
step: 190, loss: 6.0780341300414875e-05
step: 200, loss: 3.6529840144794434e-05
step: 210, loss: 2.865797432605177e-05
step: 220, loss: 2.4832181225065142e-05
step: 230, loss: 4.4991171307628974e-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=0.9887133182844244, f1=0.9819413092550789,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029528182931244373
step: 10, loss: 0.0001282725133933127
step: 20, loss: 4.039449413539842e-05
step: 30, loss: 3.8845621020300314e-05
step: 40, loss: 3.296357681392692e-05
step: 50, loss: 8.485796570312232e-05
step: 60, loss: 2.915676486736629e-05
step: 70, loss: 2.004532507271506e-05
step: 80, loss: 0.0003509772359393537
step: 90, loss: 0.0004034497251268476
step: 100, loss: 2.383750688750297e-05
step: 110, loss: 0.0004879927437286824
step: 120, loss: 0.003447391325607896
step: 130, loss: 0.0004572449834086001
step: 140, loss: 2.7793814297183417e-05
step: 150, loss: 3.621195355663076e-05
step: 160, loss: 5.5886041081976146e-05
step: 170, loss: 2.5394567273906432e-05
step: 180, loss: 3.595375528675504e-05
step: 190, loss: 3.091743201366626e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.00011064630234614015

```

```

step: 210, loss: 1.8451097275828943e-05
step: 220, loss: 2.5532397557981312e-05
step: 230, loss: 2.5547225959599018e-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=0.9887640449438202, f1=0.9808773903262092,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.379663237661589e-05
step: 10, loss: 0.0003631869039963931
step: 20, loss: 2.8694921638816595e-05
step: 30, loss: 5.7561399444239214e-05
step: 40, loss: 0.014794508926570415
step: 50, loss: 2.5718834876897745e-05
step: 60, loss: 0.00020934206258971244
step: 70, loss: 2.8687385565717705e-05
step: 80, loss: 0.00032430357532575727
step: 90, loss: 0.00027307349955663085
step: 100, loss: 3.560939148883335e-05
step: 110, loss: 4.250729034538381e-05
step: 120, loss: 1.7285106878262013e-05
step: 130, loss: 2.2049656763556413e-05
step: 140, loss: 3.212548472220078e-05
step: 150, loss: 2.393448266957421e-05
step: 160, loss: 0.0006091526011005044
step: 170, loss: 2.336451507289894e-05
step: 180, loss: 8.667818474350497e-05
step: 190, loss: 2.416919232928194e-05
step: 200, loss: 1.9259447071817704e-05
step: 210, loss: 0.0008642731700092554
step: 220, loss: 2.4302937163156457e-05
step: 230, loss: 5.8005454775411636e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9887640449438202, f1=0.9808773903262092,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.7981697965296917e-05
step: 10, loss: 1.48674798765569e-05
step: 20, loss: 3.822763392236084e-05
step: 30, loss: 2.632218638609629e-05
step: 40, loss: 3.3574473491171375e-05

```

```

step: 50, loss: 0.0003216204058844596
step: 60, loss: 4.12742629123386e-05
step: 70, loss: 2.6020497898571193e-05
step: 80, loss: 0.00024894048692658544
step: 90, loss: 7.549954170826823e-05
step: 100, loss: 3.8038113416405395e-05
step: 110, loss: 1.8749124137684703e-05
step: 120, loss: 2.7544152544578537e-05
step: 130, loss: 2.5301285859313793e-05
step: 140, loss: 2.0369496269267984e-05
step: 150, loss: 4.267732583684847e-05
step: 160, loss: 0.00010927527182502672
step: 170, loss: 2.4038217816269025e-05
step: 180, loss: 0.00015386224549729377
step: 190, loss: 0.0001668165496084839
step: 200, loss: 0.00019971969595644623
step: 210, loss: 2.3666296328883618e-05
step: 220, loss: 4.281330620869994e-05
step: 230, loss: 2.441506876493804e-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=0.9887640449438202, f1=0.9819819819819819,
best_f1=0.9841628959276018
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:12, 195.83it/s]
load_f1 = 0.9898305084745763
real_f1 = 0.9898305084745763
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 238.99it/s]

```

1.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

Downloading: 100% 433/433 [00:00<00:00, 398kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.75MB/s]
Downloading: 100% 440M/440M [00:10<00:00, 43.3MB/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: 0.5872337222099304
step: 10, loss: 0.562247097492218
step: 20, loss: 0.5235984325408936
step: 30, loss: 0.12509267032146454
step: 40, loss: 0.1216445043683052
step: 50, loss: 0.2712486982345581
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.10238832235336304
step: 70, loss: 0.2576923966407776

step: 80, loss: 0.041017964482307434
step: 90, loss: 0.6799295544624329
step: 100, loss: 0.11475103348493576
step: 110, loss: 0.06727404147386551
step: 120, loss: 0.13917037844657898
step: 130, loss: 0.1299179047346115
step: 140, loss: 0.07692074775695801
step: 150, loss: 0.08179348707199097
step: 160, loss: 0.041857052594423294
step: 170, loss: 0.21621514856815338
step: 180, loss: 0.09907803684473038
step: 190, loss: 0.004791893530637026
step: 200, loss: 0.1387174427509308
step: 210, loss: 0.06090255454182625
step: 220, loss: 0.2841092050075531
step: 230, loss: 0.20333383977413177
step: 240, loss: 0.08258482813835144
step: 250, loss: 0.026919441297650337
step: 260, loss: 0.1368577927350998
step: 270, loss: 0.008510288782417774
step: 280, loss: 0.035654496401548386
step: 290, loss: 0.1095036119222641
step: 300, loss: 0.08214574307203293
step: 310, loss: 0.11171992123126984
step: 320, loss: 0.13072654604911804
step: 330, loss: 0.03461281582713127
step: 340, loss: 0.14298786222934723
step: 350, loss: 0.05358462035655975
step: 360, loss: 0.021831875666975975
step: 370, loss: 0.21577972173690796
step: 380, loss: 0.037763357162475586
step: 390, loss: 0.12842640280723572
step: 400, loss: 0.30116844177246094
step: 410, loss: 0.12020035088062286
step: 420, loss: 0.057971879839897156
step: 430, loss: 0.1567964255809784
step: 440, loss: 0.046623777598142624
step: 450, loss: 0.008130006492137909
step: 460, loss: 0.008695781230926514
step: 470, loss: 0.19729149341583252
step: 480, loss: 0.06590892374515533
step: 490, loss: 0.06595151871442795
step: 500, loss: 0.04721258953213692
step: 510, loss: 0.0813099667429924
step: 520, loss: 0.055376920849084854
step: 530, loss: 0.004300418309867382
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.92201199815413, f1=0.9136822773186409,
best_f1=0.9136822773186409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07243433594703674
step: 10, loss: 0.058443307876586914
step: 20, loss: 0.008794976398348808
step: 30, loss: 0.022878829389810562
step: 40, loss: 0.037129782140254974
step: 50, loss: 0.15549640357494354
step: 60, loss: 0.13428810238838196
step: 70, loss: 0.038679201155900955
step: 80, loss: 0.09197419136762619
step: 90, loss: 0.08323237299919128
step: 100, loss: 0.06134437397122383
step: 110, loss: 0.019912729039788246
step: 120, loss: 0.08137855678796768
step: 130, loss: 0.08979734778404236
step: 140, loss: 0.025968793779611588
step: 150, loss: 0.11064473539590836
step: 160, loss: 0.038875117897987366
step: 170, loss: 0.03213677927851677
step: 180, loss: 0.032356347888708115
step: 190, loss: 0.04240201413631439
step: 200, loss: 0.020954877138137817
step: 210, loss: 0.10304740816354752
step: 220, loss: 0.07295770943164825
step: 230, loss: 0.004659734666347504
step: 240, loss: 0.09627924114465714
step: 250, loss: 0.007899772375822067
step: 260, loss: 0.0054758330807089806
step: 270, loss: 0.2421984076499939
step: 280, loss: 0.08992019295692444
step: 290, loss: 0.07029059529304504
step: 300, loss: 0.08677898347377777
step: 310, loss: 0.012040549889206886
step: 320, loss: 0.18024089932441711
step: 330, loss: 0.03161190450191498
step: 340, loss: 0.01979500986635685
step: 350, loss: 0.002509815851226449
step: 360, loss: 0.16569791734218597
step: 370, loss: 0.15901191532611847

```

```

step: 380, loss: 0.05431158468127251
step: 390, loss: 0.07417327165603638
step: 400, loss: 0.12664194405078888
step: 410, loss: 0.09019561111927032
step: 420, loss: 0.029261650517582893
step: 430, loss: 0.020647361874580383
step: 440, loss: 0.12300138920545578
step: 450, loss: 0.01054761465638876
step: 460, loss: 0.0171829704195261
step: 470, loss: 0.1849687397480011
step: 480, loss: 0.3324280381202698
step: 490, loss: 0.020449737086892128
step: 500, loss: 0.15230391919612885
step: 510, loss: 0.01791049726307392
step: 520, loss: 0.07984356582164764
step: 530, loss: 0.09390762448310852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9279026217228464, f1=0.920486435921422,
best_f1=0.920486435921422
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08437766134738922
step: 10, loss: 0.012025410309433937
step: 20, loss: 0.011195804923772812
step: 30, loss: 0.08570071309804916
step: 40, loss: 0.005675991531461477
step: 50, loss: 0.13597548007965088
step: 60, loss: 0.0034425915218889713
step: 70, loss: 0.006170362234115601
step: 80, loss: 0.00994893629103899
step: 90, loss: 0.004218655172735453
step: 100, loss: 0.06837465614080429
step: 110, loss: 0.021207699552178383
step: 120, loss: 0.07067082822322845
step: 130, loss: 0.006201490759849548
step: 140, loss: 0.12181615084409714
step: 150, loss: 0.051229365170001984
step: 160, loss: 0.06516372412443161
step: 170, loss: 0.0441213883459568
step: 180, loss: 0.032735783606767654
step: 190, loss: 0.005722671747207642

```

```

step: 200, loss: 0.07310470938682556
step: 210, loss: 0.16797851026058197
step: 220, loss: 0.10428943485021591
step: 230, loss: 0.10964766889810562
step: 240, loss: 0.03262338787317276
step: 250, loss: 0.12420999258756638
step: 260, loss: 0.008577442727982998
step: 270, loss: 0.03304964676499367
step: 280, loss: 0.07989399135112762
step: 290, loss: 0.00916663371026516
step: 300, loss: 0.010037213563919067
step: 310, loss: 0.05188683420419693
step: 320, loss: 0.019738823175430298
step: 330, loss: 0.003181347157806158
step: 340, loss: 0.0428011529147625
step: 350, loss: 0.06424305588006973
step: 360, loss: 0.03530039265751839
step: 370, loss: 0.07213643938302994
step: 380, loss: 0.027049947530031204
step: 390, loss: 0.027586333453655243
step: 400, loss: 0.046596016734838486
step: 410, loss: 0.05232035741209984
step: 420, loss: 0.04866361245512962
step: 430, loss: 0.004521740134805441
step: 440, loss: 0.03806234523653984
step: 450, loss: 0.10366497933864594
step: 460, loss: 0.053108904510736465
step: 470, loss: 0.11272463202476501
step: 480, loss: 0.01588636450469494
step: 490, loss: 0.012598000466823578
step: 500, loss: 0.029229890555143356
step: 510, loss: 0.11235358566045761
step: 520, loss: 0.036753613501787186
step: 530, loss: 0.02733941562473774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9320843091334895, f1=0.9227889564810482,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009302067570388317
step: 10, loss: 0.004803432617336512

```

step: 20, loss: 0.003083182731643319
step: 30, loss: 0.0015267775161191821
step: 40, loss: 0.0018732500029727817
step: 50, loss: 0.02828027307987213
step: 60, loss: 0.011662346310913563
step: 70, loss: 0.00936957448720932
step: 80, loss: 0.0009353895438835025
step: 90, loss: 0.07541661709547043
step: 100, loss: 0.0038988736923784018
step: 110, loss: 0.024095110595226288
step: 120, loss: 0.00041914897155947983
step: 130, loss: 0.003236740594729781
step: 140, loss: 0.0023153459187597036
step: 150, loss: 0.003931831568479538
step: 160, loss: 0.04334872215986252
step: 170, loss: 0.00622062012553215
step: 180, loss: 0.005279418081045151
step: 190, loss: 0.029763907194137573
step: 200, loss: 0.006601117551326752
step: 210, loss: 0.019193705171346664
step: 220, loss: 0.0016628012526780367
step: 230, loss: 0.010541831143200397
step: 240, loss: 0.01874619349837303
step: 250, loss: 0.01756545528769493
step: 260, loss: 0.0009611653513275087
step: 270, loss: 0.0029603152070194483
step: 280, loss: 0.09926515072584152
step: 290, loss: 0.013057323172688484
step: 300, loss: 0.00607810216024518
step: 310, loss: 0.008139475248754025
step: 320, loss: 0.004844526294618845
step: 330, loss: 0.018086325377225876
step: 340, loss: 0.004157870076596737
step: 350, loss: 0.004392082337290049
step: 360, loss: 0.1537913680076599
step: 370, loss: 0.004598650150001049
step: 380, loss: 0.012174966745078564
step: 390, loss: 0.01674581691622734
step: 400, loss: 0.004406021907925606
step: 410, loss: 0.01574896275997162
step: 420, loss: 0.001973688369616866
step: 430, loss: 0.08834104239940643
step: 440, loss: 0.010938228107988834
step: 450, loss: 0.006627824157476425
step: 460, loss: 0.00647340202704072
step: 470, loss: 0.03206278756260872
step: 480, loss: 0.034195274114608765
step: 490, loss: 0.0035636844113469124

```
step: 500, loss: 0.09118322283029556
step: 510, loss: 0.02080003172159195
step: 520, loss: 0.08043699711561203
step: 530, loss: 0.008834950625896454
/usr/local/lib/python3.7/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=0.9178789300797747, f1=0.916274694261524,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029059430584311485
step: 10, loss: 0.06661726534366608
step: 20, loss: 0.00772786233574152
step: 30, loss: 0.008359567262232304
step: 40, loss: 0.009005092084407806
step: 50, loss: 0.0021404416766017675
step: 60, loss: 0.018494846299290657
step: 70, loss: 0.0009654473979026079
step: 80, loss: 0.0002561491564847529
step: 90, loss: 0.0006886174087412655
step: 100, loss: 0.014841745607554913
step: 110, loss: 0.0014818140771239996
step: 120, loss: 0.015345031395554543
step: 130, loss: 0.0007740847649984062
step: 140, loss: 0.002000275067985058
step: 150, loss: 0.019883230328559875
step: 160, loss: 0.0003271649475209415
step: 170, loss: 0.13726861774921417
step: 180, loss: 0.003991129342466593
step: 190, loss: 0.007688224781304598
step: 200, loss: 0.012468477711081505
step: 210, loss: 0.00229054456576705
step: 220, loss: 0.004040545783936977
step: 230, loss: 0.04735272005200386
step: 240, loss: 0.001935226027853787
step: 250, loss: 0.015573562122881413
step: 260, loss: 0.025506330654025078
step: 270, loss: 0.0017004155088216066
step: 280, loss: 0.009180421940982342
step: 290, loss: 0.17160126566886902
step: 300, loss: 0.004062921740114689
step: 310, loss: 0.01223693322390318
step: 320, loss: 0.011614874005317688
step: 330, loss: 0.016468603163957596
step: 340, loss: 0.00900956429541111
step: 350, loss: 0.0010999729856848717
```

```

step: 360, loss: 0.01919548027217388
step: 370, loss: 0.03348761051893234
step: 380, loss: 0.0006706639542244375
step: 390, loss: 0.001384381321258843
step: 400, loss: 0.004131756257265806
step: 410, loss: 0.0013671477790921926
step: 420, loss: 0.001176500809378922
step: 430, loss: 0.003968208562582731
step: 440, loss: 0.012893587350845337
step: 450, loss: 0.011450204066932201
step: 460, loss: 0.0004926351248286664
step: 470, loss: 0.00969952903687954
step: 480, loss: 0.06570644676685333
step: 490, loss: 0.0009255704353563488
step: 500, loss: 0.01775391399860382
step: 510, loss: 0.0034192304592579603
step: 520, loss: 0.004537784028798342
step: 530, loss: 0.024175139144062996
/usr/local/lib/python3.7/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=0.9252736792003807, f1=0.9070986183897094,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003641781397163868
step: 10, loss: 0.00020474701886996627
step: 20, loss: 0.09904660284519196
step: 30, loss: 0.0005487282760441303
step: 40, loss: 0.019928406924009323
step: 50, loss: 0.0134828956797719
step: 60, loss: 0.0003854139940813184
step: 70, loss: 0.000222171816858463
step: 80, loss: 0.0032235688995569944
step: 90, loss: 0.0021879347041249275
step: 100, loss: 0.01797710731625557
step: 110, loss: 0.010363617911934853
step: 120, loss: 0.003670202335342765
step: 130, loss: 0.00032548545277677476
step: 140, loss: 0.08871591091156006
step: 150, loss: 0.0007280012359842658
step: 160, loss: 0.002071262104436755
step: 170, loss: 0.003532449249178171
step: 180, loss: 0.002116824733093381
step: 190, loss: 0.0018939293222501874
step: 200, loss: 0.002426150254905224
step: 210, loss: 0.027200322598218918

```

```

step: 220, loss: 0.0021433888468891382
step: 230, loss: 0.0015984164783731103
step: 240, loss: 0.08784101903438568
step: 250, loss: 0.0002274761936860159
step: 260, loss: 0.001219524652697146
step: 270, loss: 0.057633232325315475
step: 280, loss: 0.002577020088210702
step: 290, loss: 0.0006494931876659393
step: 300, loss: 0.005187612026929855
step: 310, loss: 0.03418124094605446
step: 320, loss: 0.003918723668903112
step: 330, loss: 0.0005716295563615859
step: 340, loss: 0.08360618352890015
step: 350, loss: 0.007446203380823135
step: 360, loss: 0.003436059458181262
step: 370, loss: 0.010352143086493015
step: 380, loss: 0.0005948430043645203
step: 390, loss: 0.00662241829559207
step: 400, loss: 0.0017503792187198997
step: 410, loss: 0.03074537217617035
step: 420, loss: 0.03932002931833267
step: 430, loss: 0.01396610215306282
step: 440, loss: 0.0011706664226949215
step: 450, loss: 0.00047523799003101885
step: 460, loss: 0.0008669861708767712
step: 470, loss: 0.003018422285094857
step: 480, loss: 0.008612841367721558
step: 490, loss: 0.0047157215885818005
step: 500, loss: 0.0005401313537731767
step: 510, loss: 0.007313009351491928
step: 520, loss: 0.0013677238021045923
step: 530, loss: 0.004885832313448191
/usr/local/lib/python3.7/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=0.9248937175247992, f1=0.9109099571224393,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011774891754612327
step: 10, loss: 0.0017161156283691525
step: 20, loss: 0.000937602249905467
step: 30, loss: 0.0003575697774067521
step: 40, loss: 0.00020761108316946775
step: 50, loss: 0.000195049520698376
step: 60, loss: 0.0029030158184468746
step: 70, loss: 0.00047848932445049286

```


step: 80, loss: 0.0033276958856731653
step: 90, loss: 0.0009084976627491415
step: 100, loss: 0.0018219061894342303
step: 110, loss: 0.00014923610433470458
step: 120, loss: 0.00035338266752660275
step: 130, loss: 0.0004760370356962085
step: 140, loss: 0.00015953672118484974
step: 150, loss: 0.00016513877199031413
step: 160, loss: 0.0004777032299898565
step: 170, loss: 0.0001234862138517201
step: 180, loss: 0.03924304246902466
step: 190, loss: 0.0020173629745841026
step: 200, loss: 0.00026935679488815367
step: 210, loss: 0.008514560759067535
step: 220, loss: 0.0008403625688515604
step: 230, loss: 0.03415752947330475
step: 240, loss: 0.000905821449123323
step: 250, loss: 0.002033023862168193
step: 260, loss: 0.0007515358156524599
step: 270, loss: 0.0002003058762056753
step: 280, loss: 0.008991394191980362
step: 290, loss: 0.00011415433255024254
step: 300, loss: 0.003866911865770817
step: 310, loss: 0.002525156131014228
step: 320, loss: 0.00030891792266629636
step: 330, loss: 0.002148380968719721
step: 340, loss: 0.022084487602114677
step: 350, loss: 0.003348212456330657
step: 360, loss: 0.0015655715251341462
step: 370, loss: 0.0003086586657445878
step: 380, loss: 0.0018237062031403184
step: 390, loss: 0.00034393215901218355
step: 400, loss: 0.02769545651972294
step: 410, loss: 0.0028648527804762125
step: 420, loss: 0.0006586028612218797
step: 430, loss: 0.0002427015861030668
step: 440, loss: 0.005182140972465277
step: 450, loss: 0.0007235835073515773
step: 460, loss: 0.0038342818152159452
step: 470, loss: 0.00023116129159461707
step: 480, loss: 0.1565122753381729
step: 490, loss: 0.0010114500764757395
step: 500, loss: 0.0019288596231490374
step: 510, loss: 0.001108529744669795
step: 520, loss: 0.010345050133764744
step: 530, loss: 0.071779765188694
/usr/local/lib/python3.7/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=0.925756186984418, f1=0.9222222222222223,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012352422345429659
step: 10, loss: 0.044468797743320465
step: 20, loss: 7.386403740383685e-05
step: 30, loss: 0.0005584176396951079
step: 40, loss: 0.0001020460476865992
step: 50, loss: 0.0022163300309330225
step: 60, loss: 0.0033700198400765657
step: 70, loss: 0.0004904621164314449
step: 80, loss: 0.004232289269566536
step: 90, loss: 0.0008282099151983857
step: 100, loss: 7.433102291543037e-05
step: 110, loss: 0.12682048976421356
step: 120, loss: 0.00047346940846182406
step: 130, loss: 0.0004782166797667742
step: 140, loss: 0.004299823194742203
step: 150, loss: 0.0123373344540596
step: 160, loss: 0.05791632831096649
step: 170, loss: 0.004642742220312357
step: 180, loss: 8.070201147347689e-05
step: 190, loss: 0.0024172975681722164
step: 200, loss: 0.0009785895235836506
step: 210, loss: 0.0020520007237792015
step: 220, loss: 0.07380320876836777
step: 230, loss: 0.0005271710688248277
step: 240, loss: 0.00016961822984740138
step: 250, loss: 0.00018224833183921874
step: 260, loss: 0.0001769998052623123
step: 270, loss: 0.0023557015229016542
step: 280, loss: 0.003779330290853977
step: 290, loss: 0.00045609293738380075
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 300, loss: 0.006632956676185131
step: 310, loss: 0.020655792206525803
step: 320, loss: 0.0036893868818879128
step: 330, loss: 0.0002477035450283438
step: 340, loss: 0.002625210676342249
step: 350, loss: 0.0012916751438751817
step: 360, loss: 0.0033422999549657106
step: 370, loss: 6.540895265061408e-05
step: 380, loss: 0.037233296781778336
step: 390, loss: 0.07879333198070526
step: 400, loss: 0.0021063508465886116

```

```

step: 410, loss: 9.402869909536093e-05
step: 420, loss: 0.002379290061071515
step: 430, loss: 0.04397496581077576
step: 440, loss: 0.012301240116357803
step: 450, loss: 0.0014100634725764394
step: 460, loss: 0.0005242509650997818
step: 470, loss: 0.0004085745895281434
step: 480, loss: 4.499522037804127e-05
step: 490, loss: 0.05622845143079758
step: 500, loss: 0.0035629193298518658
step: 510, loss: 0.02848917618393898
step: 520, loss: 0.010031981393694878
step: 530, loss: 0.0036545160692185163
/usr/local/lib/python3.7/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=0.9276164130935916, f1=0.9240801117838845,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014326549135148525
step: 10, loss: 0.00010501452925382182
step: 20, loss: 0.017081139609217644
step: 30, loss: 0.0008816898334771395
step: 40, loss: 0.00012148715177318081
step: 50, loss: 0.013598714955151081
step: 60, loss: 0.00011106029705842957
step: 70, loss: 2.3051639800542034e-05
step: 80, loss: 0.0007551027229055762
step: 90, loss: 0.00018773390911519527
step: 100, loss: 3.8800761103630066e-05
step: 110, loss: 0.0001755584089551121
step: 120, loss: 0.0002801111259032041
step: 130, loss: 0.00025787949562072754
step: 140, loss: 0.00050546188140288
step: 150, loss: 8.243149204645306e-05
step: 160, loss: 0.0008312153513543308
step: 170, loss: 0.014133873395621777
step: 180, loss: 0.002280057407915592
step: 190, loss: 4.053608063259162e-05
step: 200, loss: 0.00042757042683660984
step: 210, loss: 2.809512079693377e-05
step: 220, loss: 0.00025085589732043445
step: 230, loss: 9.575425792718306e-05
step: 240, loss: 4.352356700110249e-05
step: 250, loss: 7.720228313701227e-05
step: 260, loss: 0.01124679483473301

```

```

step: 270, loss: 2.7167869120603427e-05
step: 280, loss: 4.869440454058349e-05
step: 290, loss: 0.14167651534080505
step: 300, loss: 0.0007546232081949711
step: 310, loss: 0.011376741342246532
step: 320, loss: 0.011804724112153053
step: 330, loss: 6.333023338811472e-05
step: 340, loss: 4.701538273366168e-05
step: 350, loss: 0.02758418209850788
step: 360, loss: 5.523906656890176e-05
step: 370, loss: 0.011180557310581207
step: 380, loss: 0.00012337286898400635
step: 390, loss: 9.834632510319352e-05
step: 400, loss: 0.0013613853370770812
step: 410, loss: 2.9398293918347917e-05
step: 420, loss: 2.79465220955899e-05
step: 430, loss: 3.104093411820941e-05
step: 440, loss: 0.003660112852230668
step: 450, loss: 4.2253785068169236e-05
step: 460, loss: 0.00013050004781689495
step: 470, loss: 2.0961922928108834e-05
step: 480, loss: 4.745604019262828e-05
step: 490, loss: 0.0025239840615540743
step: 500, loss: 0.001227866392582655
step: 510, loss: 0.00015147600788623095
step: 520, loss: 2.3394455638481304e-05
step: 530, loss: 0.00045550227514468133
/usr/local/lib/python3.7/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=0.9281045751633987, f1=0.916588566073102,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.868987449910492e-05
step: 10, loss: 4.8073688958538696e-05
step: 20, loss: 2.2734986487193964e-05
step: 30, loss: 2.5718887627590448e-05
step: 40, loss: 0.00023570637858938426
step: 50, loss: 0.0001927571283886209
step: 60, loss: 0.00012504142068792135
step: 70, loss: 0.00010936996841337532
step: 80, loss: 2.3308684831135906e-05
step: 90, loss: 0.00017286656657233834
step: 100, loss: 2.6411664293846115e-05
step: 110, loss: 6.055078483768739e-05
step: 120, loss: 4.6826680772937834e-05

```

```

step: 130, loss: 0.00011199726577615365
step: 140, loss: 0.004240436479449272
step: 150, loss: 2.9518030714825727e-05
step: 160, loss: 3.530651520122774e-05
step: 170, loss: 0.0157045628875494
step: 180, loss: 0.00011773745063692331
step: 190, loss: 0.002631689887493849
step: 200, loss: 0.00010159672820009291
step: 210, loss: 0.00026339988107793033
step: 220, loss: 0.04160144552588463
step: 230, loss: 0.00020500559185165912
step: 240, loss: 0.00020453977049328387
step: 250, loss: 9.926678467309102e-05
step: 260, loss: 0.014463327825069427
step: 270, loss: 0.00042042971472255886
step: 280, loss: 0.026093676686286926
step: 290, loss: 0.00035343095078133047
step: 300, loss: 2.2484924556920305e-05
step: 310, loss: 0.002637322759255767
step: 320, loss: 0.0024701254442334175
step: 330, loss: 0.0001906196994241327
step: 340, loss: 3.575603477656841e-05
step: 350, loss: 9.873395174508914e-05
step: 360, loss: 0.00010893334547290578
step: 370, loss: 0.0014512479538097978
step: 380, loss: 0.0009370660409331322
step: 390, loss: 0.0010818983428180218
step: 400, loss: 0.0024658844340592623
step: 410, loss: 0.000933610019274056
step: 420, loss: 0.006696183700114489
step: 430, loss: 3.3968168281717226e-05
step: 440, loss: 0.0012638490879908204
step: 450, loss: 2.2775810066377744e-05
step: 460, loss: 3.770055263885297e-05
step: 470, loss: 0.0013251476921141148
step: 480, loss: 7.760692096780986e-05
step: 490, loss: 0.0007842586492188275
step: 500, loss: 0.001125585287809372
step: 510, loss: 7.617196388309821e-05
step: 520, loss: 5.218785008764826e-05
step: 530, loss: 0.0002460963441990316
/usr/local/lib/python3.7/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=0.9166269652215341, f1=0.9102196752626551,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002299302868777886
step: 10, loss: 0.0036421457771211863
step: 20, loss: 7.901016942923889e-05
step: 30, loss: 0.00025991996517404914
step: 40, loss: 0.0014154160162433982
step: 50, loss: 0.0005466886796057224
step: 60, loss: 0.004336371552199125
step: 70, loss: 0.00017711505643092096
step: 80, loss: 4.554925180855207e-05
step: 90, loss: 6.430311623262241e-05
step: 100, loss: 0.0004399821045808494
step: 110, loss: 9.803044667933136e-05
step: 120, loss: 0.00020821804355364293
step: 130, loss: 3.322433985886164e-05
step: 140, loss: 0.0028989636339247227
step: 150, loss: 2.4962719180621207e-05
step: 160, loss: 4.491902654990554e-05
step: 170, loss: 4.799326052307151e-05
step: 180, loss: 0.00015371212793979794
step: 190, loss: 0.00013520190259441733
step: 200, loss: 0.0016463055508211255
step: 210, loss: 4.424090002430603e-05
step: 220, loss: 0.0009561249753460288
step: 230, loss: 0.2906455993652344
step: 240, loss: 0.0001907560945255682
step: 250, loss: 6.079388185753487e-05
step: 260, loss: 0.00012700243678409606
step: 270, loss: 0.0004281824512872845
step: 280, loss: 8.566249744035304e-05
step: 290, loss: 3.17378289764747e-05
step: 300, loss: 0.0001551858731545508
step: 310, loss: 4.426549276104197e-05
step: 320, loss: 0.00016311062790919095
step: 330, loss: 4.891870776191354e-05
step: 340, loss: 2.2299202100839466e-05
step: 350, loss: 0.00032973141060210764
step: 360, loss: 1.94233962247381e-05
step: 370, loss: 0.00043732605990953743
step: 380, loss: 7.555847696494311e-05
step: 390, loss: 2.2105303287389688e-05
step: 400, loss: 0.00022932776482775807
step: 410, loss: 4.030377749586478e-05
step: 420, loss: 8.575587708037347e-05
step: 430, loss: 1.767255707818549e-05
step: 440, loss: 0.00028957347967661917
step: 450, loss: 0.00021935928089078516
step: 460, loss: 0.0005443893023766577
```

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 470, loss: 3.0162458642735146e-05
step: 480, loss: 0.0002571014338172972
step: 490, loss: 0.00019312025688122958
step: 500, loss: 0.0002016183134401217
step: 510, loss: 0.0001052635780069977
step: 520, loss: 3.1826275517232716e-05
step: 530, loss: 2.2582216843147762e-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=0.9245810055865922, f1=0.9201485608170844,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.688478161871899e-05
step: 10, loss: 2.292858152941335e-05
step: 20, loss: 2.6150664780288935e-05
step: 30, loss: 0.00015624152729287744
step: 40, loss: 4.2569565266603604e-05
step: 50, loss: 0.0035102174151688814
step: 60, loss: 0.0012484448961913586
step: 70, loss: 6.250416481634602e-05
step: 80, loss: 4.1845167288556695e-05
step: 90, loss: 0.0046163881197571754
step: 100, loss: 6.586040399270132e-05
step: 110, loss: 0.0002687636879272759
step: 120, loss: 1.5340581740019843e-05
step: 130, loss: 0.0003776243538595736
step: 140, loss: 0.0006018269923515618
step: 150, loss: 0.00011085822188761085
step: 160, loss: 1.853296453191433e-05
step: 170, loss: 1.8004078810918145e-05
step: 180, loss: 3.98106494685635e-05
step: 190, loss: 0.00012169677938800305
step: 200, loss: 0.00011290319525869563
step: 210, loss: 2.775543907773681e-05
step: 220, loss: 2.4038921765168197e-05
step: 230, loss: 5.7808676501736045e-05
step: 240, loss: 2.0443882021936588e-05
step: 250, loss: 1.9050879927817732e-05
step: 260, loss: 0.0008062243578024209
step: 270, loss: 4.211103077977896e-05
step: 280, loss: 0.0004890226991847157
step: 290, loss: 7.68331010476686e-05
step: 300, loss: 0.007024080026894808
step: 310, loss: 0.00012509783846326172

```

```

step: 320, loss: 3.1886058422969654e-05
step: 330, loss: 5.4725693189539015e-05
step: 340, loss: 1.9710198102984577e-05
step: 350, loss: 5.0718619604595006e-05
step: 360, loss: 0.003343293908983469
step: 370, loss: 4.1309882362838835e-05
step: 380, loss: 3.061929965042509e-05
step: 390, loss: 0.001049391576088965
step: 400, loss: 7.921731594251469e-05
step: 410, loss: 3.0537525162799284e-05
step: 420, loss: 0.008529266342520714
step: 430, loss: 0.004084838088601828
step: 440, loss: 1.6696534657967277e-05
step: 450, loss: 2.3062755644787103e-05
step: 460, loss: 8.426594285992905e-05
step: 470, loss: 1.762779720593244e-05
step: 480, loss: 0.00011444473057053983
step: 490, loss: 0.00023073704505804926
step: 500, loss: 0.0001238993281731382
step: 510, loss: 0.00090501963859424
step: 520, loss: 2.0425391994649544e-05
step: 530, loss: 1.743037319101859e-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=0.9278734295020938, f1=0.9208430913348946,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017973966896533966
step: 10, loss: 1.6547530321986414e-05
step: 20, loss: 1.3790900993626565e-05
step: 30, loss: 0.001691011362709105
step: 40, loss: 1.765015986165963e-05
step: 50, loss: 0.0064614019356667995
step: 60, loss: 1.981805507966783e-05
step: 70, loss: 1.4532188288285397e-05
step: 80, loss: 1.6275615053018555e-05
step: 90, loss: 0.0033012880012392998
step: 100, loss: 2.8604676117538475e-05
step: 110, loss: 6.32573792245239e-05
step: 120, loss: 1.81343039002968e-05
step: 130, loss: 2.8033524358761497e-05
step: 140, loss: 8.991893264465034e-05
step: 150, loss: 0.0007386620854958892
step: 160, loss: 0.007400709204375744
step: 170, loss: 0.00022371431987266988

```



```

step: 180, loss: 1.4304977412393782e-05
step: 190, loss: 2.5148538043140434e-05
step: 200, loss: 1.95051670743851e-05
step: 210, loss: 2.0298484741942957e-05
step: 220, loss: 1.5251026525220368e-05
step: 230, loss: 1.2781360965163913e-05
step: 240, loss: 8.971260831458494e-05
step: 250, loss: 1.385795167152537e-05
step: 260, loss: 0.0007739966968074441
step: 270, loss: 1.3507708899851423e-05
step: 280, loss: 1.4338511391542852e-05
step: 290, loss: 1.4204374565451872e-05
step: 300, loss: 1.709133357508108e-05
step: 310, loss: 0.00024037431285250932
step: 320, loss: 0.00016571072046644986
step: 330, loss: 5.7118893892038614e-05
step: 340, loss: 0.0004537277272902429
step: 350, loss: 1.4182046470523346e-05
step: 360, loss: 1.4148512491374277e-05
step: 370, loss: 0.0006249244324862957
step: 380, loss: 2.8351068976917304e-05
step: 390, loss: 0.0006224028766155243
step: 400, loss: 0.0004951730370521545
step: 410, loss: 0.0002024844870902598
step: 420, loss: 1.5586412700940855e-05
step: 430, loss: 7.791751704644412e-05
step: 440, loss: 4.346921559772454e-05
step: 450, loss: 0.0001665365998633206
step: 460, loss: 0.00011256598372710869
step: 470, loss: 5.232978946878575e-05
step: 480, loss: 1.4949398064345587e-05
step: 490, loss: 1.7158337868750095e-05
step: 500, loss: 1.3533841411117464e-05
step: 510, loss: 1.7340909835183993e-05
step: 520, loss: 1.4547093087458052e-05
step: 530, loss: 1.2662165318033658e-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=0.9238673517048108, f1=0.9210526315789475,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000805013463832438
step: 10, loss: 1.1309897672617808e-05
step: 20, loss: 3.0483113732771017e-05
step: 30, loss: 1.6823123587528244e-05

```

step: 40, loss: 1.7050384485628456e-05
step: 50, loss: 0.00010842637857422233
step: 60, loss: 1.2677043741859961e-05
step: 70, loss: 7.436883606715128e-05
step: 80, loss: 1.6320153008564375e-05
step: 90, loss: 1.9221670299884863e-05
step: 100, loss: 0.0004485231766011566
step: 110, loss: 1.2736663848045282e-05
step: 120, loss: 0.01697619818150997
step: 130, loss: 0.005551123060286045
step: 140, loss: 0.0001848790270742029
step: 150, loss: 1.101183624996338e-05
step: 160, loss: 8.457733201794326e-05
step: 170, loss: 7.863244536565617e-05
step: 180, loss: 3.0606322980020195e-05
step: 190, loss: 0.0026933629997074604
step: 200, loss: 2.8525890229502693e-05
step: 210, loss: 1.836137926147785e-05
step: 220, loss: 1.3381118151301052e-05
step: 230, loss: 0.00015817703388165683
step: 240, loss: 1.2609991244971752e-05
step: 250, loss: 2.087597022182308e-05
step: 260, loss: 6.842240691184998e-05
step: 270, loss: 0.00041158252861350775
step: 280, loss: 1.1827693015220575e-05
step: 290, loss: 0.0015434365486726165
step: 300, loss: 7.875078881625086e-05
step: 310, loss: 6.272702739806846e-05
step: 320, loss: 0.0006638690829277039
step: 330, loss: 7.331900269491598e-05
step: 340, loss: 1.6208501619985327e-05
step: 350, loss: 3.3080468710977584e-05
step: 360, loss: 0.043619874864816666
step: 370, loss: 1.2591361155500636e-05
step: 380, loss: 3.357378955115564e-05
step: 390, loss: 0.0010277667315676808
step: 400, loss: 0.0008104584412649274
step: 410, loss: 4.379260280984454e-05
step: 420, loss: 1.5821078704902902e-05
step: 430, loss: 0.00012985292414668947
step: 440, loss: 3.014624053321313e-05
step: 450, loss: 7.777363498462364e-05
step: 460, loss: 0.00021521237795241177
step: 470, loss: 1.4513343558064662e-05
step: 480, loss: 1.4722143532708287e-05
step: 490, loss: 8.998972771223634e-05
step: 500, loss: 3.324630597489886e-05
step: 510, loss: 5.094723746879026e-05

```

step: 520, loss: 1.9191711544408463e-05
step: 530, loss: 0.0005009107990190387
/usr/local/lib/python3.7/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=0.9250474383301709, f1=0.9203791469194313,
best_f1=0.9227889564810482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.007312135305256e-05
step: 10, loss: 1.7638500139582902e-05
step: 20, loss: 1.1160886060679331e-05
step: 30, loss: 1.5504159819101915e-05
step: 40, loss: 0.04783258214592934
step: 50, loss: 0.0008585966425016522
step: 60, loss: 9.763924936123658e-06
step: 70, loss: 0.00028877094155177474
step: 80, loss: 1.5072243513714056e-05
step: 90, loss: 1.2200216588098556e-05
step: 100, loss: 0.03269917145371437
step: 110, loss: 0.0008812924497760832
step: 120, loss: 0.0002378914796281606
step: 130, loss: 0.02910585142672062
step: 140, loss: 1.4096320228418335e-05
step: 150, loss: 1.4599197129427921e-05
step: 160, loss: 1.2360388609522488e-05
step: 170, loss: 1.351511491520796e-05
step: 180, loss: 1.0073119483422488e-05
step: 190, loss: 1.3641891200677492e-05
step: 200, loss: 3.028395803994499e-05
step: 210, loss: 0.00018147760420106351
step: 220, loss: 1.0982082130794879e-05
step: 230, loss: 0.026884254068136215
step: 240, loss: 1.4386819202627521e-05
step: 250, loss: 1.4781662684981711e-05
step: 260, loss: 1.2509398402471561e-05
step: 270, loss: 1.4334679690364283e-05
step: 280, loss: 7.668999751331285e-05
step: 290, loss: 9.342978046333883e-06
step: 300, loss: 1.4628787539550103e-05
step: 310, loss: 1.578367482579779e-05
step: 320, loss: 9.92783134279307e-06
step: 330, loss: 2.686279003683012e-05
step: 340, loss: 1.6967809642665088e-05
step: 350, loss: 9.003976629173849e-06
step: 360, loss: 0.0003863342572003603
step: 370, loss: 9.670783583715092e-06

```

```

step: 380, loss: 1.816765325202141e-05
step: 390, loss: 1.5008985428721644e-05
step: 400, loss: 1.1943183380935807e-05
step: 410, loss: 1.1689859093166888e-05
step: 420, loss: 3.240087244194001e-05
step: 430, loss: 2.349467649764847e-05
step: 440, loss: 3.270901652285829e-05
step: 450, loss: 9.711772690934595e-06
step: 460, loss: 1.2367846466077026e-05
step: 470, loss: 0.007156465668231249
step: 480, loss: 9.819808838074096e-06
step: 490, loss: 1.862873978097923e-05
step: 500, loss: 4.202139098197222e-05
step: 510, loss: 1.5958778021740727e-05
step: 520, loss: 2.263011447212193e-05
step: 530, loss: 4.6765104343649e-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=0.9285380663241476, f1=0.923581809657759,
best_f1=0.9227889564810482
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, 254.76it/s]
load_f1 = 0.9304063521718824
real_f1 = 0.9293023255813954
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:17, 256.38it/s]

```

1.1.4 Amazon-Google - Running the matcher

```

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

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

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

```

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

```

Defaults for this optimization level are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5598686933517456
step: 10, loss: 0.39016976952552795
step: 20, loss: 0.3770029544830322
step: 30, loss: 0.3105994164943695
step: 40, loss: 0.17733626067638397
step: 50, loss: 0.4168606698513031
step: 60, loss: 0.3090994954109192
step: 70, loss: 0.185141459107399
step: 80, loss: 0.20895150303840637
step: 90, loss: 0.40002357959747314
step: 100, loss: 0.3846823573112488
step: 110, loss: 0.2387569695711136
step: 120, loss: 0.22027701139450073
step: 130, loss: 0.24651862680912018

```

```

step: 140, loss: 0.2253374457359314
step: 150, loss: 0.26762914657592773
step: 160, loss: 0.33078211545944214
step: 170, loss: 0.23438753187656403
step: 180, loss: 0.14480583369731903
step: 190, loss: 0.15955743193626404
step: 200, loss: 0.29255375266075134
step: 210, loss: 0.20028340816497803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5252100840336135, f1=0.5498891352549888,
best_f1=0.5498891352549888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09663771837949753
step: 10, loss: 0.23787683248519897
step: 20, loss: 0.21504586935043335
step: 30, loss: 0.1346592903137207
step: 40, loss: 0.22120410203933716
step: 50, loss: 0.15166392922401428
step: 60, loss: 0.34754031896591187
step: 70, loss: 0.09778030216693878
step: 80, loss: 0.19169017672538757
step: 90, loss: 0.0940244123339653
step: 100, loss: 0.03794458881020546
step: 110, loss: 0.15061108767986298
step: 120, loss: 0.14286278188228607
step: 130, loss: 0.07960248738527298
step: 140, loss: 0.21991153061389923
step: 150, loss: 0.22991308569908142
step: 160, loss: 0.14030534029006958
step: 170, loss: 0.14162208139896393
step: 180, loss: 0.21158161759376526
step: 190, loss: 0.23311640322208405
step: 200, loss: 0.09357328712940216
step: 210, loss: 0.10331054031848907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5742574257425742, f1=0.5914893617021276,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030683482065796852
step: 10, loss: 0.23342949151992798
step: 20, loss: 0.14095504581928253
step: 30, loss: 0.22777880728244781
step: 40, loss: 0.10715016722679138
step: 50, loss: 0.1305210143327713
step: 60, loss: 0.16946904361248016
step: 70, loss: 0.10143307596445084
step: 80, loss: 0.2347191423177719
step: 90, loss: 0.17338448762893677
step: 100, loss: 0.15907926857471466
step: 110, loss: 0.13047517836093903
step: 120, loss: 0.08418386429548264
step: 130, loss: 0.15123796463012695
step: 140, loss: 0.14696379005908966
step: 150, loss: 0.1488105058670044
step: 160, loss: 0.07233249396085739
step: 170, loss: 0.18515363335609436
step: 180, loss: 0.07974705845117569
step: 190, loss: 0.19422675669193268
step: 200, loss: 0.10927093029022217
step: 210, loss: 0.15666598081588745
/usr/local/lib/python3.7/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=0.5707762557077626, f1=0.5124378109452736,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41017791628837585
step: 10, loss: 0.04975622147321701
step: 20, loss: 0.09580319374799728
step: 30, loss: 0.058733269572257996
step: 40, loss: 0.014554197899997234
step: 50, loss: 0.1696815937757492
step: 60, loss: 0.2411274015903473
step: 70, loss: 0.15198948979377747
step: 80, loss: 0.2784309983253479
step: 90, loss: 0.1298123598098755
step: 100, loss: 0.24111305177211761
step: 110, loss: 0.158572256565094

```



```

step: 120, loss: 0.1401047557592392
step: 130, loss: 0.08365467935800552
step: 140, loss: 0.3033895194530487
step: 150, loss: 0.0941009670495987
step: 160, loss: 0.06957479566335678
step: 170, loss: 0.06669964641332626
step: 180, loss: 0.31417426466941833
step: 190, loss: 0.06449203938245773
step: 200, loss: 0.08933966606855392
step: 210, loss: 0.24839359521865845
/usr/local/lib/python3.7/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=0.5609756097560975, f1=0.5434782608695652,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2673129439353943
step: 10, loss: 0.06349720805883408
step: 20, loss: 0.3121974468231201
step: 30, loss: 0.06477543711662292
step: 40, loss: 0.06890593469142914
step: 50, loss: 0.10543937981128693
step: 60, loss: 0.025363218039274216
step: 70, loss: 0.17449243366718292
step: 80, loss: 0.06112121045589447
step: 90, loss: 0.2655889689922333
step: 100, loss: 0.03077453374862671
step: 110, loss: 0.22072190046310425
step: 120, loss: 0.12845204770565033
step: 130, loss: 0.1124722883105278
step: 140, loss: 0.09212558716535568
step: 150, loss: 0.06517943739891052
step: 160, loss: 0.13653837144374847
step: 170, loss: 0.029008733108639717
step: 180, loss: 0.1005859375
step: 190, loss: 0.013239945285022259
step: 200, loss: 0.05534236505627632
step: 210, loss: 0.054709091782569885
/usr/local/lib/python3.7/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=0.5617977528089888, f1=0.5415019762845851,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.057444486767053604
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.03274349495768547
step: 20, loss: 0.019167179241776466
step: 30, loss: 0.006971164606511593
step: 40, loss: 0.0743473470211029
step: 50, loss: 0.011269599199295044
step: 60, loss: 0.025287549942731857
step: 70, loss: 0.006462696939706802
step: 80, loss: 0.03385995700955391
step: 90, loss: 0.18682315945625305
step: 100, loss: 0.005399980582296848
step: 110, loss: 0.03255244717001915
step: 120, loss: 0.06526565551757812
step: 130, loss: 0.10589173436164856
step: 140, loss: 0.047098591923713684
step: 150, loss: 0.015521584078669548
step: 160, loss: 0.02221456915140152
step: 170, loss: 0.18328529596328735
step: 180, loss: 0.03762838616967201
step: 190, loss: 0.1322505623102188
step: 200, loss: 0.03485976159572601
step: 210, loss: 0.04395372048020363
/usr/local/lib/python3.7/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=0.532319391634981, f1=0.5301204819277109,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05193611606955528
step: 10, loss: 0.03918690234422684
step: 20, loss: 0.008444568142294884
step: 30, loss: 0.0932883769273758
step: 40, loss: 0.013670049607753754
step: 50, loss: 0.07981932163238525
step: 60, loss: 0.025049548596143723
step: 70, loss: 0.006105369888246059
step: 80, loss: 0.23203414678573608
step: 90, loss: 0.18535643815994263
step: 100, loss: 0.01353457011282444
step: 110, loss: 0.027640823274850845
step: 120, loss: 0.10064425319433212
step: 130, loss: 0.06148994341492653
step: 140, loss: 0.015233032405376434
step: 150, loss: 0.007620153948664665
step: 160, loss: 0.014765611849725246

```

```

step: 170, loss: 0.06397248804569244
step: 180, loss: 0.06198025494813919
step: 190, loss: 0.0783245787024498
step: 200, loss: 0.010858042165637016
step: 210, loss: 0.045346904546022415
/usr/local/lib/python3.7/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=0.5306122448979592, f1=0.5183585313174947,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040432289242744446
step: 10, loss: 0.08306480944156647
step: 20, loss: 0.015737654641270638
step: 30, loss: 0.1749444305896759
step: 40, loss: 0.05998462811112404
step: 50, loss: 0.029135240241885185
step: 60, loss: 0.22300909459590912
step: 70, loss: 0.0341915525496006
step: 80, loss: 0.07771939784288406
step: 90, loss: 0.026609357446432114
step: 100, loss: 0.014723784290254116
step: 110, loss: 0.04018188267946243
step: 120, loss: 0.007097156252712011
step: 130, loss: 0.040807146579027176
step: 140, loss: 0.033219531178474426
step: 150, loss: 0.11366557329893112
step: 160, loss: 0.06208962947130203
step: 170, loss: 0.020181117579340935
step: 180, loss: 0.08734020590782166
step: 190, loss: 0.036193475127220154
step: 200, loss: 0.01671677455306053
step: 210, loss: 0.18231981992721558
/usr/local/lib/python3.7/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=0.5548098434004475, f1=0.5227817745803357,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007359281647950411
step: 10, loss: 0.02370813861489296
step: 20, loss: 0.006017932668328285
step: 30, loss: 0.007680686190724373
step: 40, loss: 0.038995008915662766

```

```

step: 50, loss: 0.09093557298183441
step: 60, loss: 0.0389908030629158
step: 70, loss: 0.08936798572540283
step: 80, loss: 0.015277009457349777
step: 90, loss: 0.13853049278259277
step: 100, loss: 0.012602911330759525
step: 110, loss: 0.021809495985507965
step: 120, loss: 0.0362430177628994
step: 130, loss: 0.002287674695253372
step: 140, loss: 0.0280891302973032
step: 150, loss: 0.04182923585176468
step: 160, loss: 0.007564271334558725
step: 170, loss: 0.015730826184153557
step: 180, loss: 0.0024418774992227554
step: 190, loss: 0.019139396026730537
step: 200, loss: 0.035888995975255966
step: 210, loss: 0.013016116805374622
/usr/local/lib/python3.7/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=0.5353159851301115, f1=0.5140562248995983,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008317526429891586
step: 10, loss: 0.05237961933016777
step: 20, loss: 0.002691912231966853
step: 30, loss: 0.0669403225183487
step: 40, loss: 0.0053863562643527985
step: 50, loss: 0.020442018285393715
step: 60, loss: 0.03719416633248329
step: 70, loss: 0.07945975661277771
step: 80, loss: 0.04134447127580643
step: 90, loss: 0.009491419419646263
step: 100, loss: 0.17108432948589325
step: 110, loss: 0.006933771073818207
step: 120, loss: 0.01120668649673462
step: 130, loss: 0.04429205134510994
step: 140, loss: 0.0343923419713974
step: 150, loss: 0.06061920151114464
step: 160, loss: 0.18754373490810394
step: 170, loss: 0.008883276022970676
step: 180, loss: 0.013052643276751041
step: 190, loss: 0.11131453514099121
step: 200, loss: 0.3001500368118286
step: 210, loss: 0.07498203963041306
/usr/local/lib/python3.7/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=0.5359712230215827, f1=0.5406427221172023,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012176276184618473
step: 10, loss: 0.04000809043645859
step: 20, loss: 0.03008384443819523
step: 30, loss: 0.002031429437920451
step: 40, loss: 0.018136685714125633
step: 50, loss: 0.0020992006175220013
step: 60, loss: 0.25499770045280457
step: 70, loss: 0.01030506007373333
step: 80, loss: 0.08130817860364914
step: 90, loss: 0.03220915049314499
step: 100, loss: 0.02958567999303341
step: 110, loss: 0.020989753305912018
step: 120, loss: 0.007643645163625479
step: 130, loss: 0.03067813068628311
step: 140, loss: 0.0045581175945699215
step: 150, loss: 0.003159858286380768
step: 160, loss: 0.004231793340295553
step: 170, loss: 0.013405407778918743
step: 180, loss: 0.00493128364905715
step: 190, loss: 0.003878149436786771
step: 200, loss: 0.00401779729872942
step: 210, loss: 0.0074470569379627705
/usr/local/lib/python3.7/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=0.5514563106796116, f1=0.537987679671458,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05238205939531326
step: 10, loss: 0.00307380105368793
step: 20, loss: 0.010595188476145267
step: 30, loss: 0.04924001917243004
step: 40, loss: 0.07484754174947739
step: 50, loss: 0.0012334446655586362
step: 60, loss: 0.010700914077460766
step: 70, loss: 0.001954845618456602
step: 80, loss: 0.020983034744858742
step: 90, loss: 0.00491857435554266
step: 100, loss: 0.029173370450735092

```

```

step: 110, loss: 0.005273088812828064
step: 120, loss: 0.013538750819861889
step: 130, loss: 0.0017924726707860827
step: 140, loss: 0.003701417939737439
step: 150, loss: 0.007149380631744862
step: 160, loss: 0.0053338767029345036
step: 170, loss: 0.001750089810229838
step: 180, loss: 0.004335935227572918
step: 190, loss: 0.020407183095812798
step: 200, loss: 0.008409718982875347
step: 210, loss: 0.0016670039622113109
/usr/local/lib/python3.7/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=0.5285412262156448, f1=0.48660714285714285,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01210823468863964
step: 10, loss: 0.0005203200853429735
step: 20, loss: 0.023618092760443687
step: 30, loss: 0.26548266410827637
step: 40, loss: 0.0027318433858454227
step: 50, loss: 0.02589152567088604
step: 60, loss: 0.0006948608788661659
step: 70, loss: 0.043159887194633484
step: 80, loss: 0.016315074637532234
step: 90, loss: 0.0009584708604961634
step: 100, loss: 0.014884503558278084
step: 110, loss: 0.0031891455873847008
step: 120, loss: 0.09157105535268784
step: 130, loss: 0.02017020806670189
step: 140, loss: 0.0010146655840799212
step: 150, loss: 0.000607042049523443
step: 160, loss: 0.006985285319387913
step: 170, loss: 0.07285317778587341
step: 180, loss: 0.022550808265805244
step: 190, loss: 0.0010946181137114763
step: 200, loss: 0.007254320662468672
step: 210, loss: 0.009718336164951324
/usr/local/lib/python3.7/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=0.534136546184739, f1=0.4946695095948827,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005621243617497385
step: 10, loss: 0.01023486815392971
step: 20, loss: 0.008865484967827797
step: 30, loss: 0.0005685767973773181
step: 40, loss: 0.0010500994976609945
step: 50, loss: 0.005096130538731813
step: 60, loss: 0.02241840399801731
step: 70, loss: 0.1326003074645996
step: 80, loss: 0.031181711703538895
step: 90, loss: 0.0017234599217772484
step: 100, loss: 0.003948406782001257
step: 110, loss: 0.016556398943066597
step: 120, loss: 0.015179336071014404
step: 130, loss: 0.0025839817244559526
step: 140, loss: 0.05396563932299614
step: 150, loss: 0.0027071682270616293
step: 160, loss: 0.0016218749806284904
step: 170, loss: 0.0044877356849610806
step: 180, loss: 0.0006885150214657187
step: 190, loss: 0.0024559549055993557
step: 200, loss: 0.011167303659021854
step: 210, loss: 0.006129298359155655
/usr/local/lib/python3.7/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=0.5424354243542435, f1=0.5297504798464492,
best_f1=0.5914893617021276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018385570729151368
step: 10, loss: 0.000630851136520505
step: 20, loss: 0.001260626595467329
step: 30, loss: 0.016938548535108566
step: 40, loss: 0.006814638618379831
step: 50, loss: 0.007466571405529976
step: 60, loss: 0.02919183112680912
step: 70, loss: 0.003323301672935486
step: 80, loss: 0.004173454362899065
step: 90, loss: 0.004244944080710411
step: 100, loss: 0.0032231814693659544
step: 110, loss: 0.10584782063961029
step: 120, loss: 0.0031826032791286707
step: 130, loss: 0.23880670964717865
step: 140, loss: 0.0015996521105989814
step: 150, loss: 0.0010233338689431548
step: 160, loss: 0.002720010234043002

```

```

step: 170, loss: 0.0008507941383868456
step: 180, loss: 0.0009153218707069755
step: 190, loss: 0.02741730399429798
step: 200, loss: 0.001972084166482091
step: 210, loss: 0.15801474452018738
/usr/local/lib/python3.7/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=0.5353728489483748, f1=0.5250501002004009,
best_f1=0.5914893617021276
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:05, 405.57it/s]
load_f1 = 0.5761467889908256
real_f1 = 0.5730994152046783
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 242.07it/s]

```


1.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5511135458946228
step: 10, loss: 0.36614981293678284
step: 20, loss: 0.2930256724357605
step: 30, loss: 0.4403647780418396
step: 40, loss: 0.4286964535713196
step: 50, loss: 0.2853653132915497
step: 60, loss: 0.28418076038360596
step: 70, loss: 0.28365567326545715
step: 80, loss: 0.2694726884365082
step: 90, loss: 0.287241667509079
step: 100, loss: 0.2560579180717468
step: 110, loss: 0.3454677164554596
step: 120, loss: 0.11085166037082672
step: 130, loss: 0.1344418227672577
step: 140, loss: 0.07381116598844528
step: 150, loss: 0.13241751492023468
step: 160, loss: 0.061889249831438065
step: 170, loss: 0.26126378774642944
step: 180, loss: 0.024352779611945152
step: 190, loss: 0.3488290309906006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6629526462395543, f1=0.6480446927374302,
best_f1=0.6480446927374302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.302174836397171
step: 10, loss: 0.1687266081571579
step: 20, loss: 0.22556905448436737
step: 30, loss: 0.11905629187822342
step: 40, loss: 0.16850240528583527
step: 50, loss: 0.08432819694280624
step: 60, loss: 0.2409660816192627
step: 70, loss: 0.17492547631263733
step: 80, loss: 0.1209685206413269
step: 90, loss: 0.12978757917881012
step: 100, loss: 0.08389350026845932
step: 110, loss: 0.1714734435081482
step: 120, loss: 0.21948076784610748
step: 130, loss: 0.0900004431605339
step: 140, loss: 0.02842702716588974
step: 150, loss: 0.08974052220582962
step: 160, loss: 0.06007702648639679
step: 170, loss: 0.1582537442445755
step: 180, loss: 0.10986218601465225
step: 190, loss: 0.13810090720653534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7795698924731183, f1=0.7947368421052631,
best_f1=0.7947368421052631
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032860368490219116
step: 10, loss: 0.14549991488456726
step: 20, loss: 0.08960597962141037
step: 30, loss: 0.02392146736383438
step: 40, loss: 0.18043825030326843
step: 50, loss: 0.1321508139371872
step: 60, loss: 0.0947292372584343
step: 70, loss: 0.059443339705467224
step: 80, loss: 0.055236347019672394
step: 90, loss: 0.059543050825595856

```

```

step: 100, loss: 0.09312521666288376
step: 110, loss: 0.015843670815229416
step: 120, loss: 0.027132930234074593
step: 130, loss: 0.013420112431049347
step: 140, loss: 0.018612325191497803
step: 150, loss: 0.12148139625787735
step: 160, loss: 0.03170456364750862
step: 170, loss: 0.09393501281738281
step: 180, loss: 0.0419924296438694
step: 190, loss: 0.07842104136943817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7862796833773087, f1=0.8083989501312335,
best_f1=0.8083989501312335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021567821502685547
step: 10, loss: 0.07963600009679794
step: 20, loss: 0.07570616900920868
step: 30, loss: 0.014059053733944893
step: 40, loss: 0.01533444318920374
step: 50, loss: 0.05935794115066528
step: 60, loss: 0.09041378647089005
step: 70, loss: 0.15455304086208344
step: 80, loss: 0.06736120581626892
step: 90, loss: 0.024894563481211662
step: 100, loss: 0.0897703617811203
step: 110, loss: 0.008761562407016754
step: 120, loss: 0.04236215725541115
step: 130, loss: 0.12205258011817932
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.008013001643121243
step: 150, loss: 0.038078419864177704
step: 160, loss: 0.007783171720802784
step: 170, loss: 0.11360546946525574
step: 180, loss: 0.004070499911904335
step: 190, loss: 0.10288755595684052
/usr/local/lib/python3.7/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=0.7654320987654321, f1=0.7355163727959698,
best_f1=0.8083989501312335

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05398346483707428
step: 10, loss: 0.006709904409945011
step: 20, loss: 0.06872320175170898
step: 30, loss: 0.020735226571559906
step: 40, loss: 0.2773658335208893
step: 50, loss: 0.08134280145168304
step: 60, loss: 0.15432527661323547
step: 70, loss: 0.03432551026344299
step: 80, loss: 0.0804884061217308
step: 90, loss: 0.04310274124145508
step: 100, loss: 0.004185428377240896
step: 110, loss: 0.11924834549427032
step: 120, loss: 0.015327166765928268
step: 130, loss: 0.07193490862846375
step: 140, loss: 0.0781630426645279
step: 150, loss: 0.09904003143310547
step: 160, loss: 0.23315685987472534
step: 170, loss: 0.036203812807798386
step: 180, loss: 0.040594808757305145
step: 190, loss: 0.08890661597251892
/usr/local/lib/python3.7/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=0.7855153203342619, f1=0.7747252747252747,
best_f1=0.8083989501312335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005720395594835281
step: 10, loss: 0.0748656764626503
step: 20, loss: 0.1093919649720192
step: 30, loss: 0.0013562669046223164
step: 40, loss: 0.09066138416528702
step: 50, loss: 0.23004648089408875
step: 60, loss: 0.010170633904635906
step: 70, loss: 0.006263600662350655
step: 80, loss: 0.043676767498254776
step: 90, loss: 0.03291713446378708
step: 100, loss: 0.0025219698436558247
step: 110, loss: 0.004514559172093868
step: 120, loss: 0.1163613349199295
step: 130, loss: 0.062247052788734436
step: 140, loss: 0.020437803119421005
step: 150, loss: 0.01058017648756504
step: 160, loss: 0.017915260046720505

```

```

step: 170, loss: 0.14139214158058167
step: 180, loss: 0.006807317957282066
step: 190, loss: 0.15667103230953217
/usr/local/lib/python3.7/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=0.7830687830687831, f1=0.7810026385224274,
best_f1=0.8083989501312335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008064177818596363
step: 10, loss: 0.007419307716190815
step: 20, loss: 0.07471861690282822
step: 30, loss: 0.006233245134353638
step: 40, loss: 0.028272481635212898
step: 50, loss: 0.022118275985121727
step: 60, loss: 0.07482971251010895
step: 70, loss: 0.002755653578788042
step: 80, loss: 0.023791946470737457
step: 90, loss: 0.027892323210835457
step: 100, loss: 0.002269639167934656
step: 110, loss: 0.014897915534675121
step: 120, loss: 0.0021414582151919603
step: 130, loss: 0.0013660227414220572
step: 140, loss: 0.004730453714728355
step: 150, loss: 0.010784519836306572
step: 160, loss: 0.1482010930776596
step: 170, loss: 0.2829379439353943
step: 180, loss: 0.0018913251115009189
step: 190, loss: 0.0012803797144442797
/usr/local/lib/python3.7/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=0.7668393782383419, f1=0.774869109947644,
best_f1=0.8083989501312335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025370478630065918
step: 10, loss: 0.25155678391456604
step: 20, loss: 0.023572267964482307
step: 30, loss: 0.08058388531208038
step: 40, loss: 0.007341885939240456
step: 50, loss: 0.01810416951775551
step: 60, loss: 0.002506467280909419
step: 70, loss: 0.032843202352523804
step: 80, loss: 0.00047784342314116657

```

```

step: 90, loss: 0.01279193814843893
step: 100, loss: 0.05142678692936897
step: 110, loss: 0.0018568589584901929
step: 120, loss: 0.007750571705400944
step: 130, loss: 0.0054554929956793785
step: 140, loss: 0.011417657136917114
step: 150, loss: 0.03142767399549484
step: 160, loss: 0.005091147031635046
step: 170, loss: 0.0037886195350438356
step: 180, loss: 0.05143178254365921
step: 190, loss: 0.002451101318001747
/usr/local/lib/python3.7/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=0.7853403141361257, f1=0.7513227513227512,
best_f1=0.8083989501312335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033762489911168814
step: 10, loss: 0.008617906831204891
step: 20, loss: 0.0024515422992408276
step: 30, loss: 0.001970411743968725
step: 40, loss: 0.005292022600769997
step: 50, loss: 0.007976409047842026
step: 60, loss: 0.001069548656232655
step: 70, loss: 0.0013272183714434505
step: 80, loss: 0.042072877287864685
step: 90, loss: 0.03237360715866089
step: 100, loss: 0.03666146472096443
step: 110, loss: 0.05217419192194939
step: 120, loss: 0.0031324666924774647
step: 130, loss: 0.007684705313295126
step: 140, loss: 0.01564077101647854
step: 150, loss: 0.0007042131037451327
step: 160, loss: 0.011448252014815807
step: 170, loss: 0.0005734330043196678
step: 180, loss: 0.009932183660566807
step: 190, loss: 0.0022855722345411777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7868852459016394, f1=0.7520435967302452,
best_f1=0.7520435967302452

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08906589448451996
step: 10, loss: 0.0004856364394072443
step: 20, loss: 0.014368642121553421
step: 30, loss: 0.00046649607247672975
step: 40, loss: 0.005526828579604626
step: 50, loss: 0.004457449074834585
step: 60, loss: 0.005684554576873779
step: 70, loss: 0.0234312042593956
step: 80, loss: 0.004314140882343054
step: 90, loss: 0.02093365229666233
step: 100, loss: 0.0008141116704791784
step: 110, loss: 0.0015213398728519678
step: 120, loss: 0.023732535541057587
step: 130, loss: 0.0006277438951656222
step: 140, loss: 0.014264336787164211
step: 150, loss: 0.010949566029012203
step: 160, loss: 0.04754112288355827
step: 170, loss: 0.0025799463037401438
step: 180, loss: 0.0014341797214001417
step: 190, loss: 0.0006363643333315849
/usr/local/lib/python3.7/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=0.7626666666666667, f1=0.7688172043010753,
best_f1=0.7520435967302452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018707837909460068
step: 10, loss: 0.0006626768154092133
step: 20, loss: 0.23076443374156952
step: 30, loss: 0.0016687721945345402
step: 40, loss: 0.0003722048713825643
step: 50, loss: 0.06045946851372719
step: 60, loss: 0.00021416698291432112
step: 70, loss: 0.000680813507642597
step: 80, loss: 0.00041210444760508835
step: 90, loss: 0.0023853257298469543
step: 100, loss: 0.002250246936455369
step: 110, loss: 0.0005717534222640097
step: 120, loss: 0.0006299273227341473
step: 130, loss: 0.0036185416392982006
step: 140, loss: 0.0009013825911097229
step: 150, loss: 0.0005740408087149262
step: 160, loss: 0.007743241731077433

```



```

step: 170, loss: 0.02555418387055397
step: 180, loss: 0.002178872935473919
step: 190, loss: 0.00312235951423645
/usr/local/lib/python3.7/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=0.7759562841530054, f1=0.7584269662921349,
best_f1=0.7520435967302452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032259858562611043
step: 10, loss: 0.00699902419000864
step: 20, loss: 0.006824836600571871
step: 30, loss: 0.0006665857508778572
step: 40, loss: 0.0008629310177639127
step: 50, loss: 0.00718812458217144
step: 60, loss: 0.00035740132443606853
step: 70, loss: 0.0003419149143155664
step: 80, loss: 0.0037659257650375366
step: 90, loss: 0.0046646567061543465
step: 100, loss: 0.0036034653894603252
step: 110, loss: 0.0005774987512268126
step: 120, loss: 0.001952738151885569
step: 130, loss: 0.00085121497977525
step: 140, loss: 0.0011480782413855195
step: 150, loss: 0.0005548004410229623
step: 160, loss: 0.0006402533035725355
step: 170, loss: 0.0019206968136131763
step: 180, loss: 0.00020300525648053735
step: 190, loss: 0.0005327186663635075
/usr/local/lib/python3.7/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=0.7641025641025642, f1=0.7648578811369509,
best_f1=0.7520435967302452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006280298810452223
step: 10, loss: 0.0010170809691771865
step: 20, loss: 0.02373403124511242
step: 30, loss: 0.0018235266907140613
step: 40, loss: 0.0018363507697358727
step: 50, loss: 0.06514640152454376
step: 60, loss: 0.0005349958082661033
step: 70, loss: 0.0025407637003809214
step: 80, loss: 0.00029681884916499257

```

```

step: 90, loss: 0.0004114369803573936
step: 100, loss: 0.00035779629251919687
step: 110, loss: 0.0037457644939422607
step: 120, loss: 0.010986769571900368
step: 130, loss: 0.00031786953331902623
step: 140, loss: 0.0009862077422440052
step: 150, loss: 0.00023066412541083992
step: 160, loss: 0.013197274878621101
step: 170, loss: 0.0002919745456892997
step: 180, loss: 0.0007293808157555759
step: 190, loss: 0.0013095538597553968
/usr/local/lib/python3.7/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=0.7729468599033816, f1=0.7414634146341464,
best_f1=0.7520435967302452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013066553510725498
step: 10, loss: 0.0004640839179046452
step: 20, loss: 0.00014490794274024665
step: 30, loss: 0.00023204743047244847
step: 40, loss: 0.021651888266205788
step: 50, loss: 0.0002562536974437535
step: 60, loss: 0.0007599603268317878
step: 70, loss: 0.00017829160788096488
step: 80, loss: 0.0005764488596469164
step: 90, loss: 0.003148721531033516
step: 100, loss: 0.014017585664987564
step: 110, loss: 0.0014718255260959268
step: 120, loss: 0.04364385828375816
step: 130, loss: 0.0008865348645485938
step: 140, loss: 0.0005704191280528903
step: 150, loss: 0.0011015321360900998
step: 160, loss: 0.0015104982303455472
step: 170, loss: 0.007320039439946413
step: 180, loss: 0.0026393572334200144
step: 190, loss: 0.0160041656345129
/usr/local/lib/python3.7/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=0.77, f1=0.745, best_f1=0.7520435967302452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005044964491389692
step: 10, loss: 0.0006745551363565028

```

```

step: 20, loss: 0.0004838015011046082
step: 30, loss: 0.00046950066462159157
step: 40, loss: 0.062303800135850906
step: 50, loss: 0.0029458184726536274
step: 60, loss: 0.0003889015060849488
step: 70, loss: 0.002671428956091404
step: 80, loss: 0.0024697277694940567
step: 90, loss: 0.0005931839696131647
step: 100, loss: 0.00026548223104327917
step: 110, loss: 0.001049741986207664
step: 120, loss: 0.0005475172656588256
step: 130, loss: 0.00034090320696122944
step: 140, loss: 0.0026911168824881315
step: 150, loss: 0.0008521293057128787
step: 160, loss: 0.08310685306787491
step: 170, loss: 0.0012477210257202387
step: 180, loss: 0.002953858580440283
step: 190, loss: 0.0014662024332210422
/usr/local/lib/python3.7/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=0.7681159420289856, f1=0.7360774818401937,
best_f1=0.7520435967302452
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, 211.12it/s]
load_f1 = 0.7174447174447175
real_f1 = 0.6987951807228916
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 238.29it/s]
```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

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

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

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

```
--checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

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

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

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

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

```
step: 0, loss: 0.6318175196647644
step: 10, loss: 0.3760833442211151
step: 20, loss: 0.3031081557273865
step: 30, loss: 0.37816429138183594
step: 40, loss: 0.27121251821517944
step: 50, loss: 0.27339309453964233
step: 60, loss: 0.21112506091594696
step: 70, loss: 0.36515700817108154
step: 80, loss: 0.36464688181877136
```

```

step: 90, loss: 0.2674309313297272
step: 100, loss: 0.20062685012817383
step: 110, loss: 0.2844129502773285
step: 120, loss: 0.23176157474517822
step: 130, loss: 0.05036976933479309
step: 140, loss: 0.20225289463996887
step: 150, loss: 0.30005815625190735
step: 160, loss: 0.17728833854198456
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.22131405770778656
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7308641975308643, f1=0.6744186046511628,
best_f1=0.6744186046511628
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09867233037948608
step: 10, loss: 0.1668378859758377
step: 20, loss: 0.09333700686693192
step: 30, loss: 0.14559583365917206
step: 40, loss: 0.10638365894556046
step: 50, loss: 0.08992646634578705
step: 60, loss: 0.10548007488250732
step: 70, loss: 0.04542343318462372
step: 80, loss: 0.18152756989002228
step: 90, loss: 0.09183375537395477
step: 100, loss: 0.10610926896333694
step: 110, loss: 0.0237403716892004
step: 120, loss: 0.0813600942492485
step: 130, loss: 0.12417858093976974
step: 140, loss: 0.2331565022468567
step: 150, loss: 0.26655468344688416
step: 160, loss: 0.09770911931991577
step: 170, loss: 0.03727874532341957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7677725118483412, f1=0.755656108597285,

```

```

best_f1=0.755656108597285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08989670872688293
step: 10, loss: 0.023147443309426308
step: 20, loss: 0.061599619686603546
step: 30, loss: 0.1526653915643692
step: 40, loss: 0.10432031005620956
step: 50, loss: 0.03676188364624977
step: 60, loss: 0.20451241731643677
step: 70, loss: 0.13838621973991394
step: 80, loss: 0.1719314604997635
step: 90, loss: 0.11958326399326324
step: 100, loss: 0.010411610826849937
step: 110, loss: 0.04426371306180954
step: 120, loss: 0.06267546117305756
step: 130, loss: 0.04878372699022293
step: 140, loss: 0.08451507240533829
step: 150, loss: 0.03322775661945343
step: 160, loss: 0.041415050625801086
step: 170, loss: 0.07841763645410538
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7828282828282828, f1=0.7707317073170732,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023707371205091476
step: 10, loss: 0.06843354552984238
step: 20, loss: 0.009158063679933548
step: 30, loss: 0.04908909276127815
step: 40, loss: 0.0029618723783642054
step: 50, loss: 0.048436012119054794
step: 60, loss: 0.038424137979745865
step: 70, loss: 0.046773094683885574
step: 80, loss: 0.09387144446372986
step: 90, loss: 0.04450619965791702
step: 100, loss: 0.18947798013687134
step: 110, loss: 0.08626751601696014
step: 120, loss: 0.025081835687160492
step: 130, loss: 0.10461409389972687

```

```

step: 140, loss: 0.026211660355329514
step: 150, loss: 0.05913802608847618
step: 160, loss: 0.1115373820066452
step: 170, loss: 0.016757782548666
/usr/local/lib/python3.7/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=0.7668161434977578, f1=0.7652173913043478,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08218513429164886
step: 10, loss: 0.08206253498792648
step: 20, loss: 0.03668218478560448
step: 30, loss: 0.04782785475254059
step: 40, loss: 0.1755436807870865
step: 50, loss: 0.03316662833094597
step: 60, loss: 0.04536353051662445
step: 70, loss: 0.1392083317041397
step: 80, loss: 0.12272012233734131
step: 90, loss: 0.07583899796009064
step: 100, loss: 0.024865012615919113
step: 110, loss: 0.05313515663146973
step: 120, loss: 0.009140473790466785
step: 130, loss: 0.031242452561855316
step: 140, loss: 0.01872062496840954
step: 150, loss: 0.04026201367378235
step: 160, loss: 0.1196068674325943
step: 170, loss: 0.020717613399028778
/usr/local/lib/python3.7/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=0.7718446601941746, f1=0.7710280373831775,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001126600313000381
step: 10, loss: 0.028733784332871437
step: 20, loss: 0.0069838762283325195
step: 30, loss: 0.015015797689557076
step: 40, loss: 0.011578892357647419
step: 50, loss: 0.06611495465040207
step: 60, loss: 0.07342983037233353
step: 70, loss: 0.047098059207201004
step: 80, loss: 0.05953454226255417
step: 90, loss: 0.11415161192417145

```



```

step: 100, loss: 0.00823925156146288
step: 110, loss: 0.0016664302675053477
step: 120, loss: 0.03358450159430504
step: 130, loss: 0.008375746197998524
step: 140, loss: 0.023337379097938538
step: 150, loss: 0.011112309992313385
step: 160, loss: 0.12291588634252548
step: 170, loss: 0.011120050214231014
/usr/local/lib/python3.7/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=0.7392405063291139, f1=0.759124087591241,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00894838199019432
step: 10, loss: 0.0010222174460068345
step: 20, loss: 0.0005143294692970812
step: 30, loss: 0.008851403370499611
step: 40, loss: 0.016851626336574554
step: 50, loss: 0.005228998139500618
step: 60, loss: 0.0034635639749467373
step: 70, loss: 0.0015045477775856853
step: 80, loss: 0.07231535762548447
step: 90, loss: 0.00033020551200024784
step: 100, loss: 0.015812551602721214
step: 110, loss: 0.007969509810209274
step: 120, loss: 0.016341986134648323
step: 130, loss: 0.07820048183202744
step: 140, loss: 0.0028322632424533367
step: 150, loss: 0.0028723576106131077
step: 160, loss: 0.00310440082103014
step: 170, loss: 0.0642060711979866
/usr/local/lib/python3.7/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=0.748051948051948, f1=0.7680798004987531,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010084843263030052
step: 10, loss: 0.011206403374671936
step: 20, loss: 0.0005504449363797903
step: 30, loss: 0.014707939699292183
step: 40, loss: 0.0005483152344822884
step: 50, loss: 0.0018037090776488185

```

```

step: 60, loss: 0.019668759778141975
step: 70, loss: 0.0002931736526079476
step: 80, loss: 0.01630253531038761
step: 90, loss: 0.0011856288183480501
step: 100, loss: 0.009831736795604229
step: 110, loss: 0.1240009218454361
step: 120, loss: 0.0016721631400287151
step: 130, loss: 0.0007877917378209531
step: 140, loss: 0.0003069985250476748
step: 150, loss: 0.017337709665298462
step: 160, loss: 0.015705114230513573
step: 170, loss: 0.0007838386809453368
/usr/local/lib/python3.7/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=0.7173913043478262, f1=0.7448979591836735,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005362995434552431
step: 10, loss: 0.022127680480480194
step: 20, loss: 0.020125823095440865
step: 30, loss: 0.010567126795649529
step: 40, loss: 0.013163222000002861
step: 50, loss: 0.00041162403067573905
step: 60, loss: 0.017996059730648994
step: 70, loss: 0.0034623262472450733
step: 80, loss: 0.07367021590471268
step: 90, loss: 0.1600414365530014
step: 100, loss: 0.05301075428724289
step: 110, loss: 0.013104936107993126
step: 120, loss: 0.016315490007400513
step: 130, loss: 0.1654028743505478
step: 140, loss: 0.0037551959976553917
step: 150, loss: 0.0013847590889781713
step: 160, loss: 0.06695353984832764
step: 170, loss: 0.017619190737605095
/usr/local/lib/python3.7/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=0.7446300715990454, f1=0.786046511627907,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052698764950037
step: 10, loss: 0.0007775906706228852

```

```

step: 20, loss: 0.069184310734272
step: 30, loss: 0.006480042822659016
step: 40, loss: 0.00010558222857071087
step: 50, loss: 0.054962195456027985
step: 60, loss: 0.0015878590056672692
step: 70, loss: 0.00013022072380408645
step: 80, loss: 0.004152595065534115
step: 90, loss: 0.10523255169391632
step: 100, loss: 0.0003075736458413303
step: 110, loss: 0.022529467940330505
step: 120, loss: 0.0004741350421682
step: 130, loss: 0.002766706980764866
step: 140, loss: 0.05084392800927162
step: 150, loss: 0.09514675289392471
step: 160, loss: 0.03946666792035103
step: 170, loss: 0.00022138241911306977
/usr/local/lib/python3.7/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=0.7244094488188977, f1=0.7568922305764412,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05170794576406479
step: 10, loss: 0.009430542588233948
step: 20, loss: 0.00011899983655894175
step: 30, loss: 7.663582073291764e-05
step: 40, loss: 0.00011832188465632498
step: 50, loss: 0.01821831427514553
step: 60, loss: 0.01997295953333378
step: 70, loss: 0.0004334365075919777
step: 80, loss: 0.00041962298564612865
step: 90, loss: 0.012111258693039417
step: 100, loss: 0.0008799241622909904
step: 110, loss: 0.03772689774632454
step: 120, loss: 0.00020527408923953772
step: 130, loss: 0.0001403194182785228
step: 140, loss: 0.02128683589398861
step: 150, loss: 0.05398178845643997
step: 160, loss: 0.013449851423501968
step: 170, loss: 0.0026653578970581293
/usr/local/lib/python3.7/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=0.7390180878552971, f1=0.7665847665847665,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006814703228883445
step: 10, loss: 0.0006600175402127206
step: 20, loss: 0.0013908405089750886
step: 30, loss: 0.00018615907174535096
step: 40, loss: 0.0010202323319390416
step: 50, loss: 8.497615635860711e-05
step: 60, loss: 0.0030423954594880342
step: 70, loss: 0.0680125281214714
step: 80, loss: 0.00011503665882628411
step: 90, loss: 0.007062805350869894
step: 100, loss: 0.00021214156004134566
step: 110, loss: 0.00034153799060732126
step: 120, loss: 0.0006318292580544949
step: 130, loss: 0.003795335767790675
step: 140, loss: 0.00030474993400275707
step: 150, loss: 0.0010782761964946985
step: 160, loss: 0.0006954005802981555
step: 170, loss: 0.00046114137512631714
/usr/local/lib/python3.7/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=0.7112299465240642, f1=0.755784061696658,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07127760350704193
step: 10, loss: 0.00016884601791389287
step: 20, loss: 0.00030357768991962075
step: 30, loss: 0.00011695031571434811
step: 40, loss: 0.0008869880111888051
step: 50, loss: 0.00015547031944151968
step: 60, loss: 0.0009794445941224694
step: 70, loss: 0.0015043909661471844
step: 80, loss: 0.00018049351638182998
step: 90, loss: 0.0002390090812696144
step: 100, loss: 0.00011029638699255884
step: 110, loss: 7.835211727069691e-05
step: 120, loss: 0.012161292135715485
step: 130, loss: 7.713834202149883e-05
step: 140, loss: 0.0003410734352655709
step: 150, loss: 0.009045814163982868
step: 160, loss: 0.00017040627426467836
step: 170, loss: 0.002466736827045679
/usr/local/lib/python3.7/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=0.741687979539642, f1=0.7669172932330827,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018592677952256054
step: 10, loss: 0.00017432637105230242
step: 20, loss: 0.0005870067980140448
step: 30, loss: 0.0006734552443958819
step: 40, loss: 0.0001579359668539837
step: 50, loss: 0.00011960080155404285
step: 60, loss: 0.0032711613457649946
step: 70, loss: 0.0006038395804353058
step: 80, loss: 0.18352796137332916
step: 90, loss: 0.0001090878649847582
step: 100, loss: 0.004923313390463591
step: 110, loss: 0.00011317073949612677
step: 120, loss: 0.02451283670961857
step: 130, loss: 0.0009236212936230004
step: 140, loss: 6.90230808686465e-05
step: 150, loss: 7.752571400487795e-05
step: 160, loss: 6.747336738044396e-05
step: 170, loss: 0.0012221502838656306
/usr/local/lib/python3.7/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=0.7563451776649747, f1=0.7669902912621359,
best_f1=0.7707317073170732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003829972120001912
step: 10, loss: 0.008267726749181747
step: 20, loss: 0.04708981513977051
step: 30, loss: 0.0005752829019911587
step: 40, loss: 0.00019127482664771378
step: 50, loss: 8.254260319517925e-05
step: 60, loss: 0.008542511612176895
step: 70, loss: 0.001458963262848556
step: 80, loss: 0.019869210198521614
step: 90, loss: 0.005825144704431295
step: 100, loss: 0.006828341633081436
step: 110, loss: 0.001402669702656567
step: 120, loss: 0.013815423473715782
step: 130, loss: 0.003554464317858219
step: 140, loss: 0.00012739349040202796
step: 150, loss: 0.00011982324940618128

```

```

step: 160, loss: 9.253302414435893e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.00020453333854675293
/usr/local/lib/python3.7/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=0.755784061696658, f1=0.7673267326732675,
best_f1=0.7707317073170732
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 [00:07, 273.35it/s]
load_f1 = 0.5544147843942505
real_f1 = 0.5265151515151515
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 240.80it/s]

```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

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

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

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

Downloading: 100% 433/433 [00:00<00:00, 558kB/s]

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

Downloading: 100% 440M/440M [00:11<00:00, 39.2MB/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: 0.6272323727607727
step: 10, loss: 0.6385806798934937
step: 20, loss: 0.4658887982368469
step: 30, loss: 0.22743049263954163
step: 40, loss: 0.15851449966430664
step: 50, loss: 0.12057666480541229
step: 60, loss: 0.18642884492874146
step: 70, loss: 0.21486212313175201
step: 80, loss: 0.07441890984773636
step: 90, loss: 0.14318391680717468
step: 100, loss: 0.007412274833768606
step: 110, loss: 0.07634761929512024
step: 120, loss: 0.018437886610627174
step: 130, loss: 0.03242601826786995
step: 140, loss: 0.0034218011423945427
step: 150, loss: 0.026629477739334106
step: 160, loss: 0.004104746971279383
step: 170, loss: 0.13692519068717957
step: 180, loss: 0.10696505755186081
step: 190, loss: 0.12359745055437088
step: 200, loss: 0.13385650515556335

```



```

step: 210, loss: 0.008874548599123955
step: 220, loss: 0.023258214816451073
step: 230, loss: 0.05387921258807182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9544950055493895, f1=0.9617117117117117,
best_f1=0.9617117117117117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05722786486148834
step: 10, loss: 0.006222542375326157
step: 20, loss: 0.18934249877929688
step: 30, loss: 0.09270238876342773
step: 40, loss: 0.04664423316717148
step: 50, loss: 0.009376595728099346
step: 60, loss: 0.023681897670030594
step: 70, loss: 0.1456463485956192
step: 80, loss: 0.004353194497525692
step: 90, loss: 0.05124802142381668
step: 100, loss: 0.06154876947402954
step: 110, loss: 0.11231456696987152
step: 120, loss: 0.12173692882061005
step: 130, loss: 0.07659444212913513
step: 140, loss: 0.0038891960866749287
step: 150, loss: 0.009212923236191273
step: 160, loss: 0.03206859901547432
step: 170, loss: 0.006207712460309267
step: 180, loss: 0.02679995819926262
step: 190, loss: 0.0127957072108984
step: 200, loss: 0.0020653177052736282
step: 210, loss: 0.0013955885078758001
step: 220, loss: 0.2021491378545761
step: 230, loss: 0.006043138913810253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9754464285714286, f1=0.9718785151856018,
best_f1=0.9718785151856018

```

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

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

```
step: 0, loss: 0.01019830722361803  
step: 10, loss: 0.0057363202795386314  
step: 20, loss: 0.015490676276385784  
step: 30, loss: 0.005173593293875456  
step: 40, loss: 0.020255718380212784  
step: 50, loss: 0.0059048887342214584  
step: 60, loss: 0.010516038164496422  
step: 70, loss: 0.0019013960845768452  
step: 80, loss: 0.0012743329862132668  
step: 90, loss: 0.07216915488243103  
step: 100, loss: 0.0006404817686416209  
step: 110, loss: 0.001454708050005138  
step: 120, loss: 0.002934189746156335  
step: 130, loss: 0.000836220511700958  
step: 140, loss: 0.003600483760237694  
step: 150, loss: 0.0057907686568796635  
step: 160, loss: 0.1372969150543213  
step: 170, loss: 0.02915395423769951  
step: 180, loss: 0.008364039473235607  
step: 190, loss: 0.012824025005102158  
step: 200, loss: 0.004419639706611633  
step: 210, loss: 0.0028109769336879253  
step: 220, loss: 0.0008703776402398944  
step: 230, loss: 0.0010808927472680807
```

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

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

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

```
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=0.9755011135857461, 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: 0.0006868878845125437  
step: 10, loss: 0.0013080076314508915  
step: 20, loss: 0.000904486165381968  
step: 30, loss: 0.0013364509213715792  
step: 40, loss: 0.0192860197275877  
step: 50, loss: 0.00036193046253174543  
step: 60, loss: 0.002537675201892853  
step: 70, loss: 0.0006392257055267692  
step: 80, loss: 0.0006990295951254666
```

```
step: 90, loss: 0.011435437947511673
step: 100, loss: 0.0006400416023097932
step: 110, loss: 0.0035114330239593983
step: 120, loss: 0.007092905696481466
step: 130, loss: 0.005673826672136784
step: 140, loss: 0.005239999387413263
step: 150, loss: 0.07949057221412659
step: 160, loss: 0.14980155229568481
step: 170, loss: 0.0035393196158111095
step: 180, loss: 0.0033339038491249084
step: 190, loss: 0.022008245810866356
step: 200, loss: 0.002051443560048938
step: 210, loss: 0.058992937207221985
step: 220, loss: 0.0006712919566780329
step: 230, loss: 0.03825074061751366
/usr/local/lib/python3.7/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=0.9742441209406495, f1=0.9774774774774775,
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: 0.0013091381406411529
step: 10, loss: 0.0008229855448007584
step: 20, loss: 0.004374862648546696
step: 30, loss: 0.0004353440017439425
step: 40, loss: 0.0011176750995218754
step: 50, loss: 0.001396744977682829
step: 60, loss: 0.17947295308113098
step: 70, loss: 0.00042917628888972104
step: 80, loss: 0.0007532804738730192
step: 90, loss: 0.002450081519782543
step: 100, loss: 0.0008540274575352669
step: 110, loss: 0.006187061313539743
step: 120, loss: 0.00031970220152288675
step: 130, loss: 0.0016512973234057426
step: 140, loss: 0.0011302046477794647
step: 150, loss: 0.002131666988134384
step: 160, loss: 0.000520218862220645
step: 170, loss: 0.004149973392486572
step: 180, loss: 0.00803207978606224
step: 190, loss: 0.12278515845537186
step: 200, loss: 0.0010623839916661382
step: 210, loss: 0.0015049861976876855
step: 220, loss: 0.0016624433919787407
step: 230, loss: 0.0006010479410178959
/usr/local/lib/python3.7/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=0.9733333333333333, f1=0.9678135405105438,
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: 0.003436524420976639
step: 10, loss: 0.0007014621514827013
step: 20, loss: 0.010746794752776623
step: 30, loss: 0.00032786891097202897
step: 40, loss: 0.0006343193235807121
step: 50, loss: 0.0009235026082023978
step: 60, loss: 0.0002914874057751149
step: 70, loss: 0.0006682883249595761
step: 80, loss: 0.03137895092368126
step: 90, loss: 0.003997448831796646
step: 100, loss: 0.02320202812552452
step: 110, loss: 0.015222234651446342
step: 120, loss: 0.0012149936519563198
step: 130, loss: 0.0014592596562579274
step: 140, loss: 0.0017179995775222778
step: 150, loss: 0.0006326601142063737
step: 160, loss: 0.01061405148357153
step: 170, loss: 0.00022857134172227234
step: 180, loss: 0.014899452216923237
step: 190, loss: 0.0512685552239418
step: 200, loss: 0.00048222392797470093
step: 210, loss: 0.0002578829589765519
step: 220, loss: 0.00019375498231966048
step: 230, loss: 0.03607169911265373
/usr/local/lib/python3.7/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=0.9721913236929923, f1=0.9732739420935412,
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: 0.009532948024570942
step: 10, loss: 0.00019974535098299384
step: 20, loss: 0.00030091742519289255
step: 30, loss: 0.0017323625506833196
step: 40, loss: 0.0003337813832331449
step: 50, loss: 0.0001943285169545561
step: 60, loss: 0.0012267319252714515
step: 70, loss: 0.00035568707971833646
step: 80, loss: 0.0003702811955008656

```

```

step: 90, loss: 0.001061328686773777
step: 100, loss: 0.0004941373481415212
step: 110, loss: 0.001140667824074626
step: 120, loss: 0.0009483007597737014
step: 130, loss: 0.0008502617129124701
step: 140, loss: 0.001493042684160173
step: 150, loss: 0.0010191855253651738
step: 160, loss: 0.006968920584768057
step: 170, loss: 0.0003539213794283569
step: 180, loss: 0.0009603811777196825
step: 190, loss: 0.0005344691453501582
step: 200, loss: 0.182531476020813
step: 210, loss: 7.822679617675021e-05
step: 220, loss: 0.0003056925779674202
step: 230, loss: 0.0008002744289115071
/usr/local/lib/python3.7/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=0.974472807991121, f1=0.9711111111111111,
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: 0.00033720716601237655
step: 10, loss: 0.00047923167585395277
step: 20, loss: 0.00043422909220680594
step: 30, loss: 0.000396596675273031
step: 40, loss: 0.0007318024290725589
step: 50, loss: 0.0005141128785908222
step: 60, loss: 0.021401667967438698
step: 70, loss: 0.0021765809506177902
step: 80, loss: 0.00039162859320640564
step: 90, loss: 9.291859896620736e-05
step: 100, loss: 0.000182778385351412
step: 110, loss: 0.00034098775358870625
step: 120, loss: 0.06986580789089203
step: 130, loss: 0.00010696431854739785
step: 140, loss: 0.0013086216058582067
step: 150, loss: 0.0015071099624037743
step: 160, loss: 0.00018495855329092592
step: 170, loss: 0.00024096628476399928
step: 180, loss: 0.0004797334549948573
step: 190, loss: 0.0005867363652214408
step: 200, loss: 0.02167506143450737
step: 210, loss: 0.09470930695533752
step: 220, loss: 0.00044945598347112536
step: 230, loss: 0.0005805523251183331
/usr/local/lib/python3.7/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=0.9710467706013363, f1=0.9742441209406495,
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: 0.00014758140605408698
step: 10, loss: 0.000762458483222872
step: 20, loss: 0.054682813584804535
step: 30, loss: 0.00015143741620704532
step: 40, loss: 0.025159163400530815
step: 50, loss: 0.0005447674193419516
step: 60, loss: 0.00015285734843928367
step: 70, loss: 0.0008053198107518256
step: 80, loss: 0.00020233573741279542
step: 90, loss: 0.00014673675468657166
step: 100, loss: 0.00024582733749412
step: 110, loss: 5.349839921109378e-05
step: 120, loss: 0.0004943953244946897
step: 130, loss: 0.00017864776600617915
step: 140, loss: 0.00015215321036521345
step: 150, loss: 0.0010956423357129097
step: 160, loss: 0.0006487272912636399
step: 170, loss: 0.0009262808598577976
step: 180, loss: 0.0006527770892716944
step: 190, loss: 0.001450885902158916
step: 200, loss: 0.09073755145072937
step: 210, loss: 0.0011757088359445333
step: 220, loss: 0.00014089877367950976
step: 230, loss: 0.00035260338336229324
/usr/local/lib/python3.7/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=0.9755011135857461, f1=0.9709821428571428,
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: 0.00026786362286657095
step: 10, loss: 0.0003182526561431587
step: 20, loss: 0.00014094641664996743
step: 30, loss: 0.00018804710998665541
step: 40, loss: 9.294907067669556e-05
step: 50, loss: 0.0001981976383831352
step: 60, loss: 0.0005084596923552454
step: 70, loss: 0.0001614742650417611
step: 80, loss: 0.00013402028707787395

```

```

step: 90, loss: 0.00045991790830157697
step: 100, loss: 0.0001476105971960351
step: 110, loss: 0.00019133422756567597
step: 120, loss: 0.00011090096813859418
step: 130, loss: 0.0001899990311358124
step: 140, loss: 0.0006240400252863765
step: 150, loss: 0.00022264562721829861
step: 160, loss: 0.0005493703065440059
step: 170, loss: 9.575619333190843e-05
step: 180, loss: 0.001218459801748395
step: 190, loss: 0.029231758788228035
step: 200, loss: 7.427950913552195e-05
step: 210, loss: 0.0016363936010748148
step: 220, loss: 0.0001444007793907076
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 230, loss: 0.00017055775970220566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776286353467561, f1=0.9708520179372198,
best_f1=0.9708520179372198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.42743577575311e-05
step: 10, loss: 0.00011502641427796334
step: 20, loss: 0.00021903148444835097
step: 30, loss: 0.006523788906633854
step: 40, loss: 0.017713483422994614
step: 50, loss: 7.93818398960866e-05
step: 60, loss: 0.00015815386723261327
step: 70, loss: 0.0001390679826727137
step: 80, loss: 5.288405736791901e-05
step: 90, loss: 9.110922110266984e-05
step: 100, loss: 0.0001421664928784594
step: 110, loss: 0.0002163573371944949
step: 120, loss: 7.091880979714915e-05
step: 130, loss: 0.0006295153289102018
step: 140, loss: 8.501949196215719e-05
step: 150, loss: 0.004219799768179655
step: 160, loss: 5.410830635810271e-05
step: 170, loss: 0.00045962227159179747
step: 180, loss: 6.214240420376882e-05
step: 190, loss: 4.052131771459244e-05

```

```

step: 200, loss: 0.0012162127532064915
step: 210, loss: 4.416377123561688e-05
step: 220, loss: 7.661757990717888e-05
step: 230, loss: 8.881156099960208e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9787234042553192, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.94651758344844e-05
step: 10, loss: 3.868903877446428e-05
step: 20, loss: 9.329359454568475e-05
step: 30, loss: 9.308448352385312e-05
step: 40, loss: 0.0004247940960340202
step: 50, loss: 0.0014758208999410272
step: 60, loss: 0.0001612178748473525
step: 70, loss: 4.4975560740567744e-05
step: 80, loss: 5.295778828440234e-05
step: 90, loss: 4.669333065976389e-05
step: 100, loss: 3.427479532547295e-05
step: 110, loss: 0.0001104996117646806
step: 120, loss: 4.243991497787647e-05
step: 130, loss: 6.290589226409793e-05
step: 140, loss: 0.00018641850329004228
step: 150, loss: 5.647996658808552e-05
step: 160, loss: 6.227409176062793e-05
step: 170, loss: 9.662740194471553e-05
step: 180, loss: 0.00010333344107493758
step: 190, loss: 4.9741302063921466e-05
step: 200, loss: 3.351831037434749e-05
step: 210, loss: 6.131573172751814e-05
step: 220, loss: 6.595544255105779e-05
step: 230, loss: 8.017026993911713e-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=0.9775784753363228, f1=0.9719416386083053,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 7.649317558389157e-05
step: 10, loss: 0.0002873015182558447
step: 20, loss: 7.963726966409013e-05
step: 30, loss: 7.519042992498726e-05
step: 40, loss: 0.0005234369891695678
step: 50, loss: 0.00024376453075092286
step: 60, loss: 8.788793638814241e-05
step: 70, loss: 4.418836033437401e-05
step: 80, loss: 8.64070825628005e-05
step: 90, loss: 0.00021822120470460504
step: 100, loss: 3.651299266493879e-05
step: 110, loss: 0.00027359131490811706
step: 120, loss: 0.00047032878501340747
step: 130, loss: 5.528688052436337e-05
step: 140, loss: 6.459734868258238e-05
step: 150, loss: 6.021814260748215e-05
step: 160, loss: 8.265238284366205e-05
step: 170, loss: 9.931538806995377e-05
step: 180, loss: 5.0828162784455344e-05
step: 190, loss: 4.9853486416395754e-05
step: 200, loss: 4.367194924270734e-05
step: 210, loss: 3.1111849239096045e-05
step: 220, loss: 4.270039062248543e-05
step: 230, loss: 3.887529965140857e-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=0.9753363228699552, f1=0.9752808988764046,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.1831134846434e-05
step: 10, loss: 0.0002117231342708692
step: 20, loss: 0.00012037333362968639
step: 30, loss: 5.988362318021245e-05
step: 40, loss: 7.062856457196176e-05
step: 50, loss: 0.0005223308107815683
step: 60, loss: 0.00011225743219256401
step: 70, loss: 4.9301386752631515e-05
step: 80, loss: 5.04372401337605e-05
step: 90, loss: 5.0473820010665804e-05
step: 100, loss: 4.486789111979306e-05
step: 110, loss: 0.0001111676829168573
step: 120, loss: 2.386328560533002e-05
step: 130, loss: 4.5046897866996005e-05
step: 140, loss: 5.553639857680537e-05
step: 150, loss: 3.0251918360590935e-05

```

```

step: 160, loss: 7.9291858128272e-05
step: 170, loss: 2.231388680229429e-05
step: 180, loss: 7.762817404000089e-05
step: 190, loss: 7.904394442448393e-05
step: 200, loss: 0.00021476755500771105
step: 210, loss: 7.482051296392456e-05
step: 220, loss: 3.477504651527852e-05
step: 230, loss: 0.011402742005884647
/usr/local/lib/python3.7/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=0.9765886287625419, f1=0.9731543624161074,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.721973942243494e-05
step: 10, loss: 2.268976459163241e-05
step: 20, loss: 4.895473466604017e-05
step: 30, loss: 0.003387375734746456
step: 40, loss: 0.0001044361197273247
step: 50, loss: 0.0004930353024974465
step: 60, loss: 5.982555012451485e-05
step: 70, loss: 0.00016425037756562233
step: 80, loss: 4.223866926622577e-05
step: 90, loss: 0.00010104808461619541
step: 100, loss: 7.35570429242216e-05
step: 110, loss: 3.138815736747347e-05
step: 120, loss: 0.0002785555552691221
step: 130, loss: 0.0030608484521508217
step: 140, loss: 5.001070530852303e-05
step: 150, loss: 9.856159158516675e-05
step: 160, loss: 2.5163570171571337e-05
step: 170, loss: 3.4766144381137565e-05
step: 180, loss: 6.801434210501611e-05
step: 190, loss: 0.0007400280446745455
step: 200, loss: 8.376876212423667e-05
step: 210, loss: 0.0002407409920124337
step: 220, loss: 0.0022616267669945955
step: 230, loss: 4.387773515190929e-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=0.9765886287625419, f1=0.9699666295884317,
best_f1=0.9764309764309763
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 [00:13, 189.44it/s]
load_f1 = 0.9765363128491621
real_f1 = 0.9754464285714286
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:20, 218.36it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--n_epochs 15 \
--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model \
--da span_shuffle \
--da attr_del \
--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
```

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

```
step: 0, loss: 0.6250107288360596
step: 10, loss: 0.5336310863494873
step: 20, loss: 0.5576497912406921
step: 30, loss: 0.25418514013290405
step: 40, loss: 0.2376287877559662
step: 50, loss: 0.21416375041007996
step: 60, loss: 0.09672262519598007
step: 70, loss: 0.13512010872364044
step: 80, loss: 0.02995406463742256
step: 90, loss: 0.2998794615268707
step: 100, loss: 0.02691008523106575
step: 110, loss: 0.16605685651302338
step: 120, loss: 0.1640739142894745
step: 130, loss: 0.12645606696605682
step: 140, loss: 0.03274572640657425
step: 150, loss: 0.031562693417072296
step: 160, loss: 0.02848323993384838
step: 170, loss: 0.17137165367603302
step: 180, loss: 0.08507341891527176
step: 190, loss: 0.022600600495934486
step: 200, loss: 0.22394131124019623
step: 210, loss: 0.07716228812932968
step: 220, loss: 0.18184639513492584
step: 230, loss: 0.21047306060791016
step: 240, loss: 0.07256125658750534
step: 250, loss: 0.02185937948524952
step: 260, loss: 0.2001318782567978
step: 270, loss: 0.014158102683722973
step: 280, loss: 0.07782170176506042
step: 290, loss: 0.08205235749483109
step: 300, loss: 0.053529296070337296
step: 310, loss: 0.23667268455028534
step: 320, loss: 0.10112227499485016
step: 330, loss: 0.03858266398310661
step: 340, loss: 0.06651676446199417
step: 350, loss: 0.05328337848186493
step: 360, loss: 0.12612323462963104
```

```

step: 370, loss: 0.16124598681926727
step: 380, loss: 0.02518547885119915
step: 390, loss: 0.1768956184387207
step: 400, loss: 0.2376360297203064
step: 410, loss: 0.07621609419584274
step: 420, loss: 0.06267932057380676
step: 430, loss: 0.1840473860502243
step: 440, loss: 0.04166756197810173
step: 450, loss: 0.014607933349907398
step: 460, loss: 0.030058907344937325
step: 470, loss: 0.09969624876976013
step: 480, loss: 0.11133234202861786
step: 490, loss: 0.07724140584468842
step: 500, loss: 0.08771775662899017
step: 510, loss: 0.06893256306648254
step: 520, loss: 0.025413891300559044
step: 530, loss: 0.015957102179527283
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.924791086350975, f1=0.9141531322505799,
best_f1=0.9141531322505799
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21875928342342377
step: 10, loss: 0.04125244915485382
step: 20, loss: 0.033916570246219635
step: 30, loss: 0.03410874307155609
step: 40, loss: 0.062335655093193054
step: 50, loss: 0.21917830407619476
step: 60, loss: 0.008884450420737267
step: 70, loss: 0.06140529364347458
step: 80, loss: 0.08026637136936188
step: 90, loss: 0.014972075819969177
step: 100, loss: 0.021111413836479187
step: 110, loss: 0.022003181278705597
step: 120, loss: 0.027946900576353073
step: 130, loss: 0.25066566467285156
step: 140, loss: 0.047909487038850784
step: 150, loss: 0.05479462444782257
step: 160, loss: 0.0109926862642169
step: 170, loss: 0.00949470978230238
step: 180, loss: 0.016415856778621674

```

```

step: 190, loss: 0.07817588001489639
step: 200, loss: 0.019249608740210533
step: 210, loss: 0.058360710740089417
step: 220, loss: 0.11968637257814407
step: 230, loss: 0.009107926860451698
step: 240, loss: 0.019009392708539963
step: 250, loss: 0.062056317925453186
step: 260, loss: 0.002682135673239827
step: 270, loss: 0.3412691056728363
step: 280, loss: 0.03584881126880646
step: 290, loss: 0.034576840698719025
step: 300, loss: 0.16214041411876678
step: 310, loss: 0.02273555099964142
step: 320, loss: 0.07128091901540756
step: 330, loss: 0.06393563747406006
step: 340, loss: 0.20883798599243164
step: 350, loss: 0.020369159057736397
step: 360, loss: 0.11233048141002655
step: 370, loss: 0.23799002170562744
step: 380, loss: 0.032107435166835785
step: 390, loss: 0.07000280916690826
step: 400, loss: 0.01792367734014988
step: 410, loss: 0.028841864317655563
step: 420, loss: 0.025420844554901123
step: 430, loss: 0.00541679048910737
step: 440, loss: 0.2521536648273468
step: 450, loss: 0.03124408610165119
step: 460, loss: 0.04308198019862175
step: 470, loss: 0.1217443123459816
step: 480, loss: 0.2805989980697632
step: 490, loss: 0.04420733079314232
step: 500, loss: 0.33018410205841064
step: 510, loss: 0.050037309527397156
step: 520, loss: 0.08995639532804489
step: 530, loss: 0.04652866721153259
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9284386617100371, f1=0.9257722452743199,
best_f1=0.9257722452743199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05059674754738808

```

step: 10, loss: 0.01004099939018488
step: 20, loss: 0.08791693300008774
step: 30, loss: 0.028131844475865364
step: 40, loss: 0.006508257240056992
step: 50, loss: 0.09758197516202927
step: 60, loss: 0.0959487333893776
step: 70, loss: 0.007024127524346113
step: 80, loss: 0.002558363601565361
step: 90, loss: 0.004551017191261053
step: 100, loss: 0.0982944443821907
step: 110, loss: 0.004076157696545124
step: 120, loss: 0.011721317656338215
step: 130, loss: 0.005247024353593588
step: 140, loss: 0.09138599783182144
step: 150, loss: 0.009714316576719284
step: 160, loss: 0.015967359766364098
step: 170, loss: 0.05674009025096893
step: 180, loss: 0.0285031758248806
step: 190, loss: 0.01737130805850029
step: 200, loss: 0.0418148934841156
step: 210, loss: 0.08057309687137604
step: 220, loss: 0.0919542908668518
step: 230, loss: 0.03676552325487137
step: 240, loss: 0.024616476148366928
step: 250, loss: 0.06415069103240967
step: 260, loss: 0.027432573959231377
step: 270, loss: 0.01675380766391754
step: 280, loss: 0.0929696336388588
step: 290, loss: 0.020510176196694374
step: 300, loss: 0.03477327898144722
step: 310, loss: 0.01379563007503748
step: 320, loss: 0.016513628885149956
step: 330, loss: 0.0029404284432530403
step: 340, loss: 0.010914426296949387
step: 350, loss: 0.0058428458869457245
step: 360, loss: 0.024484697729349136
step: 370, loss: 0.017326517030596733
step: 380, loss: 0.019499635323882103
step: 390, loss: 0.00348537671379745
step: 400, loss: 0.0819459855556488
step: 410, loss: 0.006964577361941338
step: 420, loss: 0.22495436668395996
step: 430, loss: 0.06625337153673172
step: 440, loss: 0.00602539861574769
step: 450, loss: 0.06651947647333145
step: 460, loss: 0.03267856687307358
step: 470, loss: 0.029248028993606567
step: 480, loss: 0.028400853276252747


```
step: 490, loss: 0.004658516496419907
step: 500, loss: 0.00720719899982214
step: 510, loss: 0.020483234897255898
step: 520, loss: 0.08089575171470642
step: 530, loss: 0.04219850152730942
/usr/local/lib/python3.7/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=0.9277673545966229, f1=0.9223573433115062,
best_f1=0.9257722452743199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011461540125310421
step: 10, loss: 0.012480171397328377
step: 20, loss: 0.004353762604296207
step: 30, loss: 0.003246728330850601
step: 40, loss: 0.08144260942935944
step: 50, loss: 0.013014134019613266
step: 60, loss: 0.0067221373319625854
step: 70, loss: 0.002066987566649914
step: 80, loss: 0.007256515324115753
step: 90, loss: 0.04588695243000984
step: 100, loss: 0.009750946424901485
step: 110, loss: 0.07578356564044952
step: 120, loss: 0.0008140677819028497
step: 130, loss: 0.01716563291847706
step: 140, loss: 0.0007300447323359549
step: 150, loss: 0.006058162078261375
step: 160, loss: 0.03085998445749283
step: 170, loss: 0.02840922586619854
step: 180, loss: 0.009406032972037792
step: 190, loss: 0.03766068071126938
step: 200, loss: 0.0028364979662001133
step: 210, loss: 0.00802785623818636
step: 220, loss: 0.011798758991062641
step: 230, loss: 0.08303429186344147
step: 240, loss: 0.030893635004758835
step: 250, loss: 0.03557414188981056
step: 260, loss: 0.03173670545220375
step: 270, loss: 0.012998091988265514
step: 280, loss: 0.15510301291942596
step: 290, loss: 0.018121743574738503
step: 300, loss: 0.0011336335446685553
step: 310, loss: 0.023109378293156624
step: 320, loss: 0.025468148291110992
step: 330, loss: 0.022458704188466072
step: 340, loss: 0.005865573883056641
```

```

step: 350, loss: 0.00867880042642355
step: 360, loss: 0.10614629834890366
step: 370, loss: 0.041620880365371704
step: 380, loss: 0.0028897621668875217
step: 390, loss: 0.005418324377387762
step: 400, loss: 0.001340636983513832
step: 410, loss: 0.046702172607183456
step: 420, loss: 0.015969812870025635
step: 430, loss: 0.06727778911590576
step: 440, loss: 0.004392447415739298
step: 450, loss: 0.010236674919724464
step: 460, loss: 0.0327238105237484
step: 470, loss: 0.0103880874812603
step: 480, loss: 0.019450584426522255
step: 490, loss: 0.009216675534844398
step: 500, loss: 0.0031230244785547256
step: 510, loss: 0.008443350903689861
step: 520, loss: 0.10093692690134048
step: 530, loss: 0.022592682391405106
/usr/local/lib/python3.7/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=0.9266266728195662, f1=0.9210406207211318,
best_f1=0.9257722452743199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009379683993756771
step: 10, loss: 0.05941794812679291
step: 20, loss: 0.003604433499276638
step: 30, loss: 0.0009393620421178639
step: 40, loss: 0.01421041414141655
step: 50, loss: 0.0023295599967241287
step: 60, loss: 0.03067232482135296
step: 70, loss: 0.010244756937026978
step: 80, loss: 0.005316277500241995
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 90, loss: 0.0698162093758583
step: 100, loss: 0.006686134729534388
step: 110, loss: 0.004509396385401487
step: 120, loss: 0.025783732533454895
step: 130, loss: 0.010336054489016533
step: 140, loss: 0.116871178150177
step: 150, loss: 0.030720608308911324
step: 160, loss: 0.0036766629200428724
step: 170, loss: 0.012197825126349926
step: 180, loss: 0.007036301773041487
step: 190, loss: 0.0008252732804976404

```

```

step: 200, loss: 0.0031238566152751446
step: 210, loss: 0.01267140544950962
step: 220, loss: 0.003300873562693596
step: 230, loss: 0.015780964866280556
step: 240, loss: 0.0029330349061638117
step: 250, loss: 0.04140329360961914
step: 260, loss: 0.02513638511300087
step: 270, loss: 0.01957814395427704
step: 280, loss: 0.013524472713470459
step: 290, loss: 0.13223378360271454
step: 300, loss: 0.12625113129615784
step: 310, loss: 0.03579799830913544
step: 320, loss: 0.22038568556308746
step: 330, loss: 0.025316039100289345
step: 340, loss: 0.03843915835022926
step: 350, loss: 0.0009549874230287969
step: 360, loss: 0.0038976543582975864
step: 370, loss: 0.006081125233322382
step: 380, loss: 0.005066091660410166
step: 390, loss: 0.0012037535198032856
step: 400, loss: 0.005785140674561262
step: 410, loss: 0.0019227720331400633
step: 420, loss: 0.04433853551745415
step: 430, loss: 0.016319192945957184
step: 440, loss: 0.013752717524766922
step: 450, loss: 0.20763060450553894
step: 460, loss: 0.21447521448135376
step: 470, loss: 0.020266348496079445
step: 480, loss: 0.02451515570282936
step: 490, loss: 0.0010664565488696098
step: 500, loss: 0.049595024436712265
step: 510, loss: 0.10214345157146454
step: 520, loss: 0.006155101116746664
step: 530, loss: 0.0007862570928409696
/usr/local/lib/python3.7/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=0.9272898961284232, f1=0.9188935771214253,
best_f1=0.9257722452743199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004043903201818466
step: 10, loss: 0.01059016864746809
step: 20, loss: 0.11898799985647202
step: 30, loss: 0.009215115569531918
step: 40, loss: 0.01044280081987381
step: 50, loss: 0.008048697374761105

```

step: 60, loss: 0.0052298009395599365
step: 70, loss: 0.0028466475196182728
step: 80, loss: 0.005311969667673111
step: 90, loss: 0.0002548668417148292
step: 100, loss: 0.002306459005922079
step: 110, loss: 0.002939398167654872
step: 120, loss: 0.08959479629993439
step: 130, loss: 0.2448611706495285
step: 140, loss: 0.04207808896899223
step: 150, loss: 0.0038147398736327887
step: 160, loss: 0.003529051085934043
step: 170, loss: 0.0037204723339527845
step: 180, loss: 0.005902436096221209
step: 190, loss: 0.03593343868851662
step: 200, loss: 0.0007720871362835169
step: 210, loss: 0.003519673366099596
step: 220, loss: 0.000975207716692239
step: 230, loss: 0.0003574024885892868
step: 240, loss: 0.01633242890238762
step: 250, loss: 0.0008645025081932545
step: 260, loss: 0.0003434877726249397
step: 270, loss: 0.08818070590496063
step: 280, loss: 0.024491125717759132
step: 290, loss: 0.0003689182340167463
step: 300, loss: 0.09860266000032425
step: 310, loss: 0.001104968599975109
step: 320, loss: 0.0018736030906438828
step: 330, loss: 0.0018683430971577764
step: 340, loss: 0.04327124357223511
step: 350, loss: 0.04018038511276245
step: 360, loss: 0.01200622227787971
step: 370, loss: 0.0049972087144851685
step: 380, loss: 0.0020740856416523457
step: 390, loss: 0.03494369611144066
step: 400, loss: 0.001217177603393793
step: 410, loss: 0.004482007585465908
step: 420, loss: 0.04383473843336105
step: 430, loss: 0.0019395425915718079
step: 440, loss: 0.0004330834199208766
step: 450, loss: 0.00022206627181731164
step: 460, loss: 0.002854064805433154
step: 470, loss: 0.004757443908601999
step: 480, loss: 0.0014879860682412982
step: 490, loss: 0.010944832116365433
step: 500, loss: 0.0021536187268793583
step: 510, loss: 0.03241245821118355
step: 520, loss: 0.003013713750988245
step: 530, loss: 0.012054461985826492

```
/usr/local/lib/python3.7/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=0.9188156638013372, f1=0.9065644465740298,
best_f1=0.9257722452743199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026516689104028046
step: 10, loss: 0.02112685516476631
step: 20, loss: 0.003690860467031598
step: 30, loss: 0.008923575282096863
step: 40, loss: 0.012304296717047691
step: 50, loss: 0.005773033946752548
step: 60, loss: 0.0035496880300343037
step: 70, loss: 0.001090831821784377
step: 80, loss: 0.043742671608924866
step: 90, loss: 0.0014936340739950538
step: 100, loss: 0.00845676101744175
step: 110, loss: 0.0006498415023088455
step: 120, loss: 0.0005231787799857557
step: 130, loss: 0.0073765176348388195
step: 140, loss: 0.00013199500972405076
step: 150, loss: 0.0002370682923356071
step: 160, loss: 0.00013515738828573376
step: 170, loss: 0.001555336988531053
step: 180, loss: 0.0017570910276845098
step: 190, loss: 0.08574385195970535
step: 200, loss: 0.0003117130836471915
step: 210, loss: 0.003648452227935195
step: 220, loss: 0.002425294602289796
step: 230, loss: 0.032147470861673355
step: 240, loss: 0.0010614757193252444
step: 250, loss: 0.001056790235452354
step: 260, loss: 0.00029749146779067814
step: 270, loss: 0.0013584636617451906
step: 280, loss: 0.0004741982847917825
step: 290, loss: 0.004388021770864725
step: 300, loss: 0.012946628965437412
step: 310, loss: 0.0003468043578322977
step: 320, loss: 0.009301493875682354
step: 330, loss: 0.0008860579109750688
step: 340, loss: 0.08735083043575287
step: 350, loss: 0.02882680855691433
step: 360, loss: 0.028775442391633987
step: 370, loss: 0.005211397539824247
step: 380, loss: 0.0010356876300647855
step: 390, loss: 0.00248689204454422
```

```
step: 400, loss: 0.027434401214122772
step: 410, loss: 0.0025534965097904205
step: 420, loss: 0.0027736250776797533
step: 430, loss: 0.001114472048357129
step: 440, loss: 0.0014923895942047238
step: 450, loss: 0.005076574627310038
step: 460, loss: 0.00117348728235811
step: 470, loss: 0.0015274628531187773
step: 480, loss: 0.15098999440670013
step: 490, loss: 0.005917516071349382
step: 500, loss: 0.00217036297544837
step: 510, loss: 0.0006151055567897856
step: 520, loss: 0.0036887198220938444
step: 530, loss: 0.015709953382611275
/usr/local/lib/python3.7/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=0.9260480452190296, f1=0.9228624535315985,
best_f1=0.9257722452743199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061941780149936676
step: 10, loss: 0.05714159458875656
step: 20, loss: 0.00042710319394245744
step: 30, loss: 0.1900741010904312
step: 40, loss: 0.08784323185682297
step: 50, loss: 0.00827659759670496
step: 60, loss: 0.015462014824151993
step: 70, loss: 0.0018962426111102104
step: 80, loss: 0.003036987967789173
step: 90, loss: 0.001319834846071899
step: 100, loss: 0.0009117678855545819
step: 110, loss: 0.0012249937281012535
step: 120, loss: 0.01696595922112465
step: 130, loss: 0.006803576834499836
step: 140, loss: 0.0054540070705115795
step: 150, loss: 0.0005413895705714822
step: 160, loss: 0.007701313588768244
step: 170, loss: 0.07417140156030655
step: 180, loss: 0.0030538386199623346
step: 190, loss: 0.007430862169712782
step: 200, loss: 0.0011200613807886839
step: 210, loss: 0.0019059760961681604
step: 220, loss: 0.0018493498209863901
step: 230, loss: 0.005334306973963976
step: 240, loss: 0.01359800435602665
step: 250, loss: 0.003110149409621954
```

```

step: 260, loss: 0.0029242797754704952
step: 270, loss: 0.0010094587923958898
step: 280, loss: 0.016353819519281387
step: 290, loss: 0.003161750268191099
step: 300, loss: 0.0002700872428249568
step: 310, loss: 0.0009710612939670682
step: 320, loss: 0.038187555968761444
step: 330, loss: 0.002534865401685238
step: 340, loss: 0.00023529907048214227
step: 350, loss: 0.00024428858887404203
step: 360, loss: 0.002339365892112255
step: 370, loss: 0.002561153843998909
step: 380, loss: 0.0008658821461722255
step: 390, loss: 0.1275612711906433
step: 400, loss: 0.0008965544984675944
step: 410, loss: 0.0023996185045689344
step: 420, loss: 0.05701284483075142
step: 430, loss: 0.06169004738330841
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.043235570192337036
step: 450, loss: 0.014666055329144001
step: 460, loss: 0.0010445354273542762
step: 470, loss: 0.007603014353662729
step: 480, loss: 0.000468417041702196
step: 490, loss: 0.023417284712195396
step: 500, loss: 0.05259472131729126
step: 510, loss: 0.004640974570065737
step: 520, loss: 0.006967483088374138
step: 530, loss: 0.005984815768897533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9328953542937587, f1=0.9272811486799445,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01988065615296364
step: 10, loss: 0.0010004739742726088
step: 20, loss: 0.016762567684054375
step: 30, loss: 0.0015287181595340371
step: 40, loss: 0.007955086417496204
step: 50, loss: 0.0010056816972792149
step: 60, loss: 0.0017144422745332122

```

step: 70, loss: 0.02000226266682148
step: 80, loss: 0.18759390711784363
step: 90, loss: 0.0001308225910179317
step: 100, loss: 0.00019758565758820623
step: 110, loss: 6.207833212101832e-05
step: 120, loss: 0.0008643838227726519
step: 130, loss: 0.13374117016792297
step: 140, loss: 0.0009577284217812121
step: 150, loss: 0.002281525870785117
step: 160, loss: 0.002916078781709075
step: 170, loss: 0.03719000518321991
step: 180, loss: 0.00030326194246299565
step: 190, loss: 0.0020193657837808132
step: 200, loss: 0.006390130612999201
step: 210, loss: 0.010601107031106949
step: 220, loss: 0.00697848992422231
step: 230, loss: 0.00010090417345054448
step: 240, loss: 0.00014192183152772486
step: 250, loss: 0.00787318404763937
step: 260, loss: 0.07086658477783203
step: 270, loss: 3.997019666712731e-05
step: 280, loss: 0.001513650524429977
step: 290, loss: 0.04752369225025177
step: 300, loss: 0.0002375897893216461
step: 310, loss: 0.0016862283227965236
step: 320, loss: 0.00011106083547929302
step: 330, loss: 0.00028506884700618684
step: 340, loss: 0.0002545973111409694
step: 350, loss: 0.0021130801178514957
step: 360, loss: 0.00014379870845004916
step: 370, loss: 0.012379368767142296
step: 380, loss: 0.004711740184575319
step: 390, loss: 0.00016388361109420657
step: 400, loss: 0.022759653627872467
step: 410, loss: 0.011005820706486702
step: 420, loss: 0.00011905279825441539
step: 430, loss: 0.0020527848973870277
step: 440, loss: 0.0016724405577406287
step: 450, loss: 0.00016613153275102377
step: 460, loss: 0.00019008411618415266
step: 470, loss: 7.498483319068328e-05
step: 480, loss: 0.00048750301357358694
step: 490, loss: 0.000773559499066323
step: 500, loss: 0.0023504195269197226
step: 510, loss: 0.0023157645482569933
step: 520, loss: 0.00032349550747312605
step: 530, loss: 0.0009741209214553237
/usr/local/lib/python3.7/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=0.9235023041474655, f1=0.9172727272727272,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038243632297962904
step: 10, loss: 0.00015031793736852705
step: 20, loss: 0.00020139366097282618
step: 30, loss: 0.12830762565135956
step: 40, loss: 0.02466922625899315
step: 50, loss: 0.000736700021661818
step: 60, loss: 0.0005058547249063849
step: 70, loss: 0.0014385840622708201
step: 80, loss: 4.036363316117786e-05
step: 90, loss: 6.734984344802797e-05
step: 100, loss: 0.002566717565059662
step: 110, loss: 0.0002940943231806159
step: 120, loss: 0.0001479332277085632
step: 130, loss: 0.006049870513379574
step: 140, loss: 0.0010345230111852288
step: 150, loss: 0.0007789410301484168
step: 160, loss: 0.0009615772287361324
step: 170, loss: 0.000313486933009699
step: 180, loss: 0.0001056750988936983
step: 190, loss: 0.00019328849157318473
step: 200, loss: 0.0006431223591789603
step: 210, loss: 0.02402842603623867
step: 220, loss: 0.0008000945090316236
step: 230, loss: 0.0029653245583176613
step: 240, loss: 0.054989468306303024
step: 250, loss: 0.003675666404888034
step: 260, loss: 0.0008339117630384862
step: 270, loss: 0.0004468917613849044
step: 280, loss: 0.00023130989575292915
step: 290, loss: 0.00016832941037137061
step: 300, loss: 3.502668914734386e-05
step: 310, loss: 3.022244527528528e-05
step: 320, loss: 0.003921724855899811
step: 330, loss: 0.012836663983762264
step: 340, loss: 0.009055448696017265
step: 350, loss: 0.17506611347198486
step: 360, loss: 0.005540866404771805
step: 370, loss: 0.009802076034247875
step: 380, loss: 0.012254969216883183
step: 390, loss: 0.0017934717470780015
step: 400, loss: 0.00136926444247365
```

```

step: 410, loss: 0.0011774327140301466
step: 420, loss: 4.337113568908535e-05
step: 430, loss: 2.2626973077422008e-05
step: 440, loss: 0.0024848785251379013
step: 450, loss: 0.0001563931436976418
step: 460, loss: 0.010198519565165043
step: 470, loss: 9.897509153233841e-05
step: 480, loss: 4.946129411109723e-05
step: 490, loss: 0.0007400544127449393
step: 500, loss: 0.00655938358977437
step: 510, loss: 0.0003621215873863548
step: 520, loss: 0.0004809077363461256
step: 530, loss: 0.00037601994699798524
/usr/local/lib/python3.7/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=0.9261311172668514, f1=0.9259427532939573,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007101847440935671
step: 10, loss: 0.004149319604039192
step: 20, loss: 0.0004203944990877062
step: 30, loss: 7.794889097567648e-05
step: 40, loss: 0.0020212430972605944
step: 50, loss: 0.010584396310150623
step: 60, loss: 0.02299872413277626
step: 70, loss: 0.002616112120449543
step: 80, loss: 0.0003056651330552995
step: 90, loss: 0.0017274419078603387
step: 100, loss: 0.0028327121399343014
step: 110, loss: 0.0003963744966313243
step: 120, loss: 0.0003296566428616643
step: 130, loss: 0.0013182489201426506
step: 140, loss: 0.0007191677577793598
step: 150, loss: 0.00015270989388227463
step: 160, loss: 0.000927078421227634
step: 170, loss: 0.0005413090693764389
step: 180, loss: 0.0029819561168551445
step: 190, loss: 0.0023234430700540543
step: 200, loss: 0.044869594275951385
step: 210, loss: 0.0010997453937307
step: 220, loss: 0.0030462152790278196
step: 230, loss: 0.00022778002312406898
step: 240, loss: 0.0011576745891943574
step: 250, loss: 7.034656300675124e-05
step: 260, loss: 0.00010000258043874055

```

```

step: 270, loss: 0.011485747061669827
step: 280, loss: 0.0001989347510971129
step: 290, loss: 0.004165291786193848
step: 300, loss: 0.0006521149189211428
step: 310, loss: 0.002312285592779517
step: 320, loss: 0.0026832956355065107
step: 330, loss: 0.0022134799510240555
step: 340, loss: 0.004702759440988302
step: 350, loss: 0.00862901471555233
step: 360, loss: 6.499871960841119e-05
step: 370, loss: 0.006257199216634035
step: 380, loss: 0.0006648687412962317
step: 390, loss: 0.00012737551878672093
step: 400, loss: 0.0006317549268715084
step: 410, loss: 0.001081216149032116
step: 420, loss: 0.0003707555588334799
step: 430, loss: 6.501408643089235e-05
step: 440, loss: 0.0006512632826343179
step: 450, loss: 0.0009865263709798455
step: 460, loss: 0.01973380520939827
step: 470, loss: 0.0008236462599597871
step: 480, loss: 0.0002969898923765868
step: 490, loss: 9.738477820064873e-05
step: 500, loss: 0.0008515906520187855
step: 510, loss: 0.00013297382974997163
step: 520, loss: 3.5183693398721516e-05
step: 530, loss: 0.0003185014065820724
/usr/local/lib/python3.7/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=0.9256820319849483, f1=0.9289667896678966,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001306844496866688
step: 10, loss: 0.03217674046754837
step: 20, loss: 0.002956575248390436
step: 30, loss: 0.00017622123414184898
step: 40, loss: 0.002177947899326682
step: 50, loss: 0.023925859481096268
step: 60, loss: 0.0038911844603717327
step: 70, loss: 0.00016021242481656373
step: 80, loss: 0.00017800339264795184
step: 90, loss: 0.001582128694280982
step: 100, loss: 0.01072614174336195
step: 110, loss: 0.020114503800868988
step: 120, loss: 0.0002851661993190646

```

```

step: 130, loss: 0.00024353861226700246
step: 140, loss: 7.213515345938504e-05
step: 150, loss: 0.0002557545667514205
step: 160, loss: 3.949810707126744e-05
step: 170, loss: 0.005246126092970371
step: 180, loss: 0.004171072971075773
step: 190, loss: 0.002702137688174844
step: 200, loss: 0.0008994857198558748
step: 210, loss: 0.0007814839482307434
step: 220, loss: 0.0021035235840827227
step: 230, loss: 0.00029727417859248817
step: 240, loss: 0.00020308559760451317
step: 250, loss: 9.617111209081486e-05
step: 260, loss: 3.9904869481688365e-05
step: 270, loss: 0.00021600966283585876
step: 280, loss: 0.0006112267728894949
step: 290, loss: 0.00151115155313164
step: 300, loss: 0.00016928551485762
step: 310, loss: 0.0006311901961453259
step: 320, loss: 0.00033162496401928365
step: 330, loss: 0.000607523019425571
step: 340, loss: 6.903434405103326e-05
step: 350, loss: 0.0011316664749756455
step: 360, loss: 0.00011216233542654663
step: 370, loss: 0.008147523738443851
step: 380, loss: 0.0010534028988331556
step: 390, loss: 0.007562472019344568
step: 400, loss: 0.0002752193540800363
step: 410, loss: 0.011224566958844662
step: 420, loss: 0.002255696104839444
step: 430, loss: 0.008961012586951256
step: 440, loss: 0.0020060832612216473
step: 450, loss: 0.0008243209449574351
step: 460, loss: 0.00574861653149128
step: 470, loss: 0.0015772365732118487
step: 480, loss: 0.0012558362213894725
step: 490, loss: 0.0026389912236481905
step: 500, loss: 0.0005214456468820572
step: 510, loss: 0.00047785695642232895
step: 520, loss: 0.00042672461131587625
step: 530, loss: 0.001506082247942686
/usr/local/lib/python3.7/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=0.9252506836827713, f1=0.9232150875617423,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010257776593789458
step: 10, loss: 0.0002520294510759413
step: 20, loss: 0.0008724513463675976
step: 30, loss: 0.007714629173278809
step: 40, loss: 0.0026181081775575876
step: 50, loss: 0.0007257011020556092
step: 60, loss: 0.00015218235785141587
step: 70, loss: 0.0058768815360963345
step: 80, loss: 0.09417713433504105
step: 90, loss: 0.036916255950927734
step: 100, loss: 0.038368891924619675
step: 110, loss: 0.00012816116213798523
step: 120, loss: 0.0006388004403561354
step: 130, loss: 0.0006500330637209117
step: 140, loss: 5.109642006573267e-05
step: 150, loss: 0.016815874725580215
step: 160, loss: 7.947490666992962e-05
step: 170, loss: 0.0003382638969924301
step: 180, loss: 0.0005284949438646436
step: 190, loss: 0.0002073692885460332
step: 200, loss: 0.005387668032199144
step: 210, loss: 0.00041554830386303365
step: 220, loss: 0.0004308974603191018
step: 230, loss: 0.0005342481890693307
step: 240, loss: 0.00017661206948105246
step: 250, loss: 0.005643775686621666
step: 260, loss: 9.013331873575225e-05
step: 270, loss: 0.00020229430811014026
step: 280, loss: 0.0001890958083095029
step: 290, loss: 0.0050109075382351875
step: 300, loss: 0.000146473670611158
step: 310, loss: 0.0006779037648811936
step: 320, loss: 0.0005706271040253341
step: 330, loss: 0.0002988605701830238
step: 340, loss: 0.001303994795307517
step: 350, loss: 0.00035873957676813006
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 360, loss: 0.0017954254290089011
step: 370, loss: 0.003334361594170332
step: 380, loss: 0.00023617272381670773
step: 390, loss: 0.001630427548661828
step: 400, loss: 0.0002188769431086257
step: 410, loss: 0.0006349982577376068
step: 420, loss: 6.727821892127395e-05
step: 430, loss: 0.023979440331459045
step: 440, loss: 0.0002132198424078524
step: 450, loss: 9.191586286760867e-05
```

```

step: 460, loss: 0.0008016889914870262
step: 470, loss: 0.0006196242175064981
step: 480, loss: 0.008547858335077763
step: 490, loss: 0.003331880085170269
step: 500, loss: 0.0001077852793969214
step: 510, loss: 0.0012801055563613772
step: 520, loss: 0.005201779305934906
step: 530, loss: 7.708552584517747e-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=0.9262371615312792, f1=0.9307479224376731,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007855987059883773
step: 10, loss: 0.0008543406729586422
step: 20, loss: 0.002253109123557806
step: 30, loss: 0.00012710278679151088
step: 40, loss: 0.0002684864157345146
step: 50, loss: 0.05794136971235275
step: 60, loss: 0.00047777523286640644
step: 70, loss: 0.000597676495090127
step: 80, loss: 0.00012410010094754398
step: 90, loss: 0.00044910484575666487
step: 100, loss: 0.0412088967859745
step: 110, loss: 0.00017004134133458138
step: 120, loss: 0.000977182760834694
step: 130, loss: 0.025816477835178375
step: 140, loss: 0.00030147534562274814
step: 150, loss: 0.0008669918170198798
step: 160, loss: 7.608409941894934e-05
step: 170, loss: 0.010063668712973595
step: 180, loss: 0.0009335193317383528
step: 190, loss: 0.00025447754887863994
step: 200, loss: 0.00025783086312003434
step: 210, loss: 9.586560918251052e-05
step: 220, loss: 7.972054299898446e-05
step: 230, loss: 0.0005856107454746962
step: 240, loss: 0.0836346447467804
step: 250, loss: 0.022995607927441597
step: 260, loss: 0.011085886508226395
step: 270, loss: 0.0007171651814132929
step: 280, loss: 5.479063474922441e-05
step: 290, loss: 0.005849149078130722
step: 300, loss: 0.003506515407934785
step: 310, loss: 0.0008135876851156354

```

```

step: 320, loss: 0.0001609502942301333
step: 330, loss: 0.0004837685846723616
step: 340, loss: 0.0039848037995398045
step: 350, loss: 0.0001297986600548029
step: 360, loss: 0.0052740261889994144
step: 370, loss: 6.220844807103276e-05
step: 380, loss: 0.00011071359040215611
step: 390, loss: 0.057929009199142456
step: 400, loss: 0.0004992802860215306
step: 410, loss: 0.015007607638835907
step: 420, loss: 0.016005191951990128
step: 430, loss: 0.00018638602341525257
step: 440, loss: 0.002779037691652775
step: 450, loss: 0.0002668591623660177
step: 460, loss: 5.747053728555329e-05
step: 470, loss: 0.00020399795903358608
step: 480, loss: 0.0057429359294474125
step: 490, loss: 7.012120477156714e-05
step: 500, loss: 0.00026295860880054533
step: 510, loss: 0.05212615802884102
step: 520, loss: 0.0001821710029616952
step: 530, loss: 0.003967090509831905
/usr/local/lib/python3.7/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=0.9280146721687299, f1=0.9253055681303757,
best_f1=0.9272811486799445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031289851176552474
step: 10, loss: 0.00490194046869874
step: 20, loss: 0.001981972251087427
step: 30, loss: 0.0017187731573358178
step: 40, loss: 0.046632248908281326
step: 50, loss: 0.00010129957809112966
step: 60, loss: 0.00012869207421317697
step: 70, loss: 0.005612397566437721
step: 80, loss: 0.0002480443217791617
step: 90, loss: 0.00017284697969444096
step: 100, loss: 0.00011093580542365089
step: 110, loss: 0.00017545628361403942
step: 120, loss: 0.002040546154603362
step: 130, loss: 0.061571765691041946
step: 140, loss: 0.00010529555584071204
step: 150, loss: 2.9446666303556412e-05
step: 160, loss: 0.00011229627853026614
step: 170, loss: 6.84762344462797e-05

```

```

step: 180, loss: 0.00022734641970600933
step: 190, loss: 0.0002488278259988874
step: 200, loss: 0.0007744913455098867
step: 210, loss: 0.0005680334870703518
step: 220, loss: 5.361156217986718e-05
step: 230, loss: 0.0022253338247537613
step: 240, loss: 0.00015838972467463464
step: 250, loss: 0.00024433020735159516
step: 260, loss: 0.0008013008628040552
step: 270, loss: 0.004021026194095612
step: 280, loss: 0.00023618480190634727
step: 290, loss: 0.00015209882985800505
step: 300, loss: 0.000218579254578799
step: 310, loss: 0.005726757925003767
step: 320, loss: 0.00022295072267297655
step: 330, loss: 0.0007981235394254327
step: 340, loss: 0.00012174806033726782
step: 350, loss: 4.3102783820359036e-05
step: 360, loss: 0.0007211176562123001
step: 370, loss: 0.0017397896153852344
step: 380, loss: 0.00015044781321194023
step: 390, loss: 0.0007865820080041885
step: 400, loss: 0.001022586366161704
step: 410, loss: 0.00016444844368379563
step: 420, loss: 0.00014685007045045495
step: 430, loss: 8.959843398770317e-05
step: 440, loss: 0.00021735849441029131
step: 450, loss: 3.8308095099637285e-05
step: 460, loss: 7.682610157644376e-05
step: 470, loss: 0.005050958599895239
step: 480, loss: 5.949040496489033e-05
step: 490, loss: 3.809733607340604e-05
step: 500, loss: 0.01592225581407547
step: 510, loss: 0.00011751625424949452
step: 520, loss: 3.7139328924240544e-05
step: 530, loss: 3.94142625737004e-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=0.927806241266884, f1=0.9292929292929294,
best_f1=0.9272811486799445
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled          : True
opt_level        : 02

```



```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:24, 233.76it/s]
load_f1 = 0.9302325581395349
real_f1 = 0.9278734295020938
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:19, 231.14it/s]

```

2 BASELINE

2.1 BASELINE STRUCTURED

2.1.1 Beer - Running the matcher

```

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

```

```

--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5484673976898193
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.42553191489361697, f1=0.3333333333333333,
best_f1=0.3333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48218342661857605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4878048780487805, f1=0.36923076923076925,
best_f1=0.36923076923076925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4577658772468567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5925925925925927, f1=0.4090909090909091,
best_f1=0.4090909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.3213229179382324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.6666666666666666, f1=0.5161290322580646,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21855618059635162
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.5833333333333334,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20412425696849823
/usr/local/lib/python3.7/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=0.35294117647058826, f1=0.5714285714285714,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34790509939193726
/usr/local/lib/python3.7/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=0.5555555555555556, f1=0.4761904761904762,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15360206365585327
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.64, f1=0.64, best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022247734013944864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.6666666666666666, f1=0.6206896551724138,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061489660292863846
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.6086956521739131,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013835974037647247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7200000000000001, f1=0.7142857142857143,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08828990161418915
/usr/local/lib/python3.7/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=0.6086956521739131, f1=0.6923076923076924,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010333404876291752
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.6153846153846153, f1=0.6923076923076924,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06039715185761452
/usr/local/lib/python3.7/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=0.6086956521739131, f1=0.6923076923076924,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008605348877608776
/usr/local/lib/python3.7/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=0.6086956521739131, f1=0.6923076923076924,
best_f1=0.7142857142857143
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, 90381.64it/s]

```

```

/usr/local/lib/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 = 0.7200000000000001
real_f1 = 0.7200000000000001
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 240.05it/s]
```

2.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6313726902008057
step: 10, loss: 0.6229407787322998
step: 20, loss: 0.41482698917388916
step: 30, loss: 0.25595155358314514
step: 40, loss: 0.21480651199817657
step: 50, loss: 0.02529362589120865
step: 60, loss: 0.0370478481054306
step: 70, loss: 0.16470833122730255
step: 80, loss: 0.07635821402072906
step: 90, loss: 0.14296014606952667
step: 100, loss: 0.0054512410424649715
step: 110, loss: 0.2137707769870758
step: 120, loss: 0.22250783443450928
step: 130, loss: 0.00855565071105957
step: 140, loss: 0.004932206589728594
step: 150, loss: 0.06164698302745819
step: 160, loss: 0.015553798526525497
step: 170, loss: 0.038609664887189865
step: 180, loss: 0.013469770550727844
step: 190, loss: 0.002342686289921403
step: 200, loss: 0.006932740565389395

```



```

step: 210, loss: 0.013503743335604668
step: 220, loss: 0.007293320260941982
step: 230, loss: 0.007103644777089357
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.9807037457434733,
best_f1=0.9807037457434733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036576322745531797
step: 10, loss: 0.006013576872646809
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.1640501618385315
step: 30, loss: 0.17808346450328827
step: 40, loss: 0.1490037739276886
step: 50, loss: 0.006575130857527256
step: 60, loss: 0.0064918105490505695
step: 70, loss: 0.1530108004808426
step: 80, loss: 0.0024336997885257006
step: 90, loss: 0.02934975177049637
step: 100, loss: 0.03926753252744675
step: 110, loss: 0.08197450637817383
step: 120, loss: 0.06417106091976166
step: 130, loss: 0.027256090193986893
step: 140, loss: 0.008578396402299404
step: 150, loss: 0.005240920931100845
step: 160, loss: 0.024421952664852142
step: 170, loss: 0.007477628532797098
step: 180, loss: 0.005570892244577408
step: 190, loss: 0.020964549854397774
step: 200, loss: 0.0034042634069919586
step: 210, loss: 0.0009279838413931429
step: 220, loss: 0.0897841677069664
step: 230, loss: 0.035311732441186905
/usr/local/lib/python3.7/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=0.9752252252252253, f1=0.9752252252252253,
best_f1=0.9807037457434733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0035862878430634737
step: 10, loss: 0.08730155229568481
step: 20, loss: 0.05709159001708031
step: 30, loss: 0.01217019371688366
step: 40, loss: 0.10653454810380936
step: 50, loss: 0.006176474504172802
step: 60, loss: 0.004381086677312851
step: 70, loss: 0.0014544507721439004
step: 80, loss: 0.000512099068146199
step: 90, loss: 0.0973469614982605
step: 100, loss: 0.00039783352985978127
step: 110, loss: 0.0019461100455373526
step: 120, loss: 0.011426118202507496
step: 130, loss: 0.0008037836523726583
step: 140, loss: 0.010913594625890255
step: 150, loss: 0.07060063630342484
step: 160, loss: 0.05223942920565605
step: 170, loss: 0.041153036057949066
step: 180, loss: 0.011639142408967018
step: 190, loss: 0.0026496557984501123
step: 200, loss: 0.0014574377564713359
step: 210, loss: 0.0025055729784071445
step: 220, loss: 0.018133092671632767
step: 230, loss: 0.0017897723009809852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876819708846584, 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: 0.0009279916994273663
step: 10, loss: 0.0002196755667682737
step: 20, loss: 0.00034604250686243176
step: 30, loss: 0.0002547982439864427
step: 40, loss: 0.03230214864015579
step: 50, loss: 0.0005457907100208104
step: 60, loss: 0.00040646755951456726
step: 70, loss: 0.00022863784397486597
step: 80, loss: 0.001004137797281146
step: 90, loss: 0.0004007723182439804
step: 100, loss: 0.00016860885079950094
step: 110, loss: 0.0007477664039470255

```

```

step: 120, loss: 0.028251010924577713
step: 130, loss: 0.0021664893720299006
step: 140, loss: 0.0006858426495455205
step: 150, loss: 0.01108548417687416
step: 160, loss: 0.0009594230796210468
step: 170, loss: 0.0011227595387026668
step: 180, loss: 0.0004427552630659193
step: 190, loss: 0.0073725017718970776
step: 200, loss: 0.001051538623869419
step: 210, loss: 0.07595030218362808
step: 220, loss: 0.0003198773192707449
step: 230, loss: 0.0210571326315403
/usr/local/lib/python3.7/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=0.9865168539325843, f1=0.9784824462061155,
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: 0.004753921180963516
step: 10, loss: 0.00023983095888979733
step: 20, loss: 0.0010386103531345725
step: 30, loss: 0.00043489784002304077
step: 40, loss: 0.00024901816505007446
step: 50, loss: 0.0001951140584424138
step: 60, loss: 0.00021169550018385053
step: 70, loss: 0.00018962944159284234
step: 80, loss: 0.000251841644058004
step: 90, loss: 0.001099374028854072
step: 100, loss: 0.0003673874889500439
step: 110, loss: 0.0019675479270517826
step: 120, loss: 7.206832378869876e-05
step: 130, loss: 0.001560223987326026
step: 140, loss: 0.0023698736913502216
step: 150, loss: 0.00015866276226006448
step: 160, loss: 0.00015490468649659306
step: 170, loss: 0.0016298199770972133
step: 180, loss: 0.00030097088892944157
step: 190, loss: 0.0001833132264437154
step: 200, loss: 0.00019995660113636404
step: 210, loss: 0.00011667318904073909
step: 220, loss: 0.0014964401489123702
step: 230, loss: 0.0001758491707732901
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9820224719101124,

```

```

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: 0.007855436764657497
step: 10, loss: 0.00016401569882873446
step: 20, loss: 0.0005073609645478427
step: 30, loss: 0.0003529949754010886
step: 40, loss: 0.0002105867606587708
step: 50, loss: 0.00022359401918947697
step: 60, loss: 8.25299066491425e-05
step: 70, loss: 0.00011207797069801018
step: 80, loss: 0.00266650621779263
step: 90, loss: 0.00029256113339215517
step: 100, loss: 0.0040875994600355625
step: 110, loss: 0.06639378517866135
step: 120, loss: 0.0002658333396539092
step: 130, loss: 0.000253872451139614
step: 140, loss: 6.778664101148024e-05
step: 150, loss: 0.00020457983191590756
step: 160, loss: 0.000554807367734611
step: 170, loss: 0.22410111129283905
step: 180, loss: 0.03842240571975708
step: 190, loss: 0.003968399949371815
step: 200, loss: 0.002726284321397543
step: 210, loss: 0.0020731547847390175
step: 220, loss: 0.004069641698151827
step: 230, loss: 0.042022813111543655
/usr/local/lib/python3.7/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=0.9744160177975528, f1=0.9752808988764046,
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: 0.007092895917594433
step: 10, loss: 0.00020104862051084638
step: 20, loss: 0.000811635167337954
step: 30, loss: 0.0004901767824776471
step: 40, loss: 0.00010026467498391867
step: 50, loss: 0.0014902532566338778
step: 60, loss: 0.003231508657336235
step: 70, loss: 0.010232103988528252
step: 80, loss: 0.0003154856967739761
step: 90, loss: 3.937800647690892e-05
step: 100, loss: 6.757522351108491e-05
step: 110, loss: 0.00039619544986635447

```

```

step: 120, loss: 0.00013641231635119766
step: 130, loss: 0.00015320109378080815
step: 140, loss: 0.007115810643881559
step: 150, loss: 0.0006940675666555762
step: 160, loss: 0.022638943046331406
step: 170, loss: 0.001933847670443356
step: 180, loss: 0.004715424962341785
step: 190, loss: 7.848426321288571e-05
step: 200, loss: 0.017798233777284622
step: 210, loss: 6.111805123509839e-05
step: 220, loss: 0.0015152664855122566
step: 230, loss: 0.00017955448129214346
/usr/local/lib/python3.7/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=0.9841986455981941, f1=0.9738339021615473,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.67193705542013e-05
step: 10, loss: 0.0025122619699686766
step: 20, loss: 5.568941560341045e-05
step: 30, loss: 8.190892549464479e-05
step: 40, loss: 0.00020092302293051034
step: 50, loss: 0.00029378151521086693
step: 60, loss: 7.254511001519859e-05
step: 70, loss: 0.0005073431530036032
step: 80, loss: 0.0004333883698564023
step: 90, loss: 4.4268126657698303e-05
step: 100, loss: 8.533514483133331e-05
step: 110, loss: 0.00017349373956676573
step: 120, loss: 0.00011877962970174849
step: 130, loss: 0.00011389509018044919
step: 140, loss: 0.00011858184006996453
step: 150, loss: 4.264447488822043e-05
step: 160, loss: 0.0003914789413101971
step: 170, loss: 6.3290775869973e-05
step: 180, loss: 8.236812573159114e-05
step: 190, loss: 9.178911568596959e-05
step: 200, loss: 7.41475960239768e-05
step: 210, loss: 0.00017674350237939507
step: 220, loss: 0.0005893924972042441
step: 230, loss: 0.00011610144429141656
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9818594104308391,

```

```

best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.677119770552963e-05
step: 10, loss: 0.054849207401275635
step: 20, loss: 0.0017184015596285462
step: 30, loss: 0.0013532806187868118
step: 40, loss: 0.014925114810466766
step: 50, loss: 3.8808542740298435e-05
step: 60, loss: 0.02826596051454544
step: 70, loss: 8.86283814907074e-05
step: 80, loss: 0.0009470906225033104
step: 90, loss: 0.18812844157218933
step: 100, loss: 0.00937458872795105
step: 110, loss: 0.00010765470506157726
step: 120, loss: 9.256033808924258e-05
step: 130, loss: 0.0007576775387860835
step: 140, loss: 0.0030474902596324682
step: 150, loss: 0.00021122170437593013
step: 160, loss: 0.03749889135360718
step: 170, loss: 0.00012378644896671176
step: 180, loss: 0.0004100118821952492
step: 190, loss: 0.0001558769727125764
step: 200, loss: 0.00019809074001386762
step: 210, loss: 0.0015410792548209429
step: 220, loss: 0.00010430542170070112
step: 230, loss: 0.00015794154023751616
/usr/local/lib/python3.7/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=0.9841986455981941, f1=0.9737742303306728,
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: 0.00031003571348264813
step: 10, loss: 9.069509542314336e-05
step: 20, loss: 0.00023412462905980647
step: 30, loss: 9.484760084887967e-05
step: 40, loss: 4.6214099711505696e-05
step: 50, loss: 0.00019961032376158983
step: 60, loss: 0.0001452641299692914
step: 70, loss: 0.00022434275888372213
step: 80, loss: 8.445489220321178e-05
step: 90, loss: 0.00015420102863572538
step: 100, loss: 9.096714347833768e-05
step: 110, loss: 0.00042143213795498013

```

```

step: 120, loss: 0.00021580870088655502
step: 130, loss: 6.096655852161348e-05
step: 140, loss: 0.0010093383025377989
step: 150, loss: 0.00020312925335019827
step: 160, loss: 3.9713057049084455e-05
step: 170, loss: 5.059472459834069e-05
step: 180, loss: 0.0001131364842876792
step: 190, loss: 0.0003442221786826849
step: 200, loss: 4.1109964513452724e-05
step: 210, loss: 4.632577474694699e-05
step: 220, loss: 5.210286326473579e-05
step: 230, loss: 0.0003669382131192833
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9798657718120806,
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: 4.176608126726933e-05
step: 10, loss: 5.1376620831433684e-05
step: 20, loss: 0.00081181077985093
step: 30, loss: 0.017944954335689545
step: 40, loss: 0.0002674549468792975
step: 50, loss: 5.527544635697268e-05
step: 60, loss: 6.021952140145004e-05
step: 70, loss: 8.970302587840706e-05
step: 80, loss: 5.89300507272128e-05
step: 90, loss: 7.670942432014272e-05
step: 100, loss: 0.00016066960233729333
step: 110, loss: 0.00039000518154352903
step: 120, loss: 3.6014516808791086e-05
step: 130, loss: 2.4571372705395333e-05
step: 140, loss: 6.108088564360514e-05
step: 150, loss: 0.0008569598430767655
step: 160, loss: 4.12961162510328e-05
step: 170, loss: 7.495753379771486e-05
step: 180, loss: 0.00011860400263685733
step: 190, loss: 4.2526808101683855e-05
step: 200, loss: 0.00011132602230645716
step: 210, loss: 3.407738768146373e-05
step: 220, loss: 4.4369771785568446e-05
step: 230, loss: 0.001070614904165268
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9798657718120806,

```

```

best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.9855127397459e-05
step: 10, loss: 3.3429656468797475e-05
step: 20, loss: 5.067562233307399e-05
step: 30, loss: 6.929194933036342e-05
step: 40, loss: 7.657520473003387e-05
step: 50, loss: 0.0026993302162736654
step: 60, loss: 0.0037782900035381317
step: 70, loss: 0.00045659055467695
step: 80, loss: 4.99307680001948e-05
step: 90, loss: 3.595140515244566e-05
step: 100, loss: 3.096348882536404e-05
step: 110, loss: 3.952340557589196e-05
step: 120, loss: 3.8189671613508835e-05
step: 130, loss: 3.578349424060434e-05
step: 140, loss: 8.250265091191977e-05
step: 150, loss: 4.044303932460025e-05
step: 160, loss: 4.445890954229981e-05
step: 170, loss: 5.427061842055991e-05
step: 180, loss: 0.00016791527741588652
step: 190, loss: 6.00349267187994e-05
step: 200, loss: 2.193399086536374e-05
step: 210, loss: 3.687510616146028e-05
step: 220, loss: 4.283936868887395e-05
step: 230, loss: 0.00012849336781073362
/usr/local/lib/python3.7/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=0.9841628959276018, f1=0.9725400457665903,
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: 0.00010504043893888593
step: 10, loss: 4.635422374121845e-05
step: 20, loss: 3.895369445672259e-05
step: 30, loss: 0.00011611251102294773
step: 40, loss: 2.3640173822059296e-05
step: 50, loss: 3.521024336805567e-05
step: 60, loss: 3.072526305913925e-05
step: 70, loss: 2.4016186216613278e-05
step: 80, loss: 3.1287727324524894e-05
step: 90, loss: 4.655759403249249e-05
step: 100, loss: 2.3371721908915788e-05
step: 110, loss: 5.6891814892878756e-05

```



```

step: 120, loss: 0.0013983440585434437
step: 130, loss: 5.0475180614739656e-05
step: 140, loss: 3.6249250115361065e-05
step: 150, loss: 4.1799408791121095e-05
step: 160, loss: 3.38683275913354e-05
step: 170, loss: 3.082197144976817e-05
step: 180, loss: 4.9737802328309044e-05
step: 190, loss: 4.6925924834795296e-05
step: 200, loss: 6.60024888929911e-05
step: 210, loss: 2.2116495529189706e-05
step: 220, loss: 0.0002912609779741615
step: 230, loss: 2.7044841772294603e-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=0.9854096520763187, f1=0.9809203142536477,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.0833212804282084e-05
step: 10, loss: 0.0008075088262557983
step: 20, loss: 3.2796298910398036e-05
step: 30, loss: 3.545235449564643e-05
step: 40, loss: 0.04915791377425194
step: 50, loss: 3.630886203609407e-05
step: 60, loss: 6.180451600812376e-05
step: 70, loss: 3.436412225710228e-05
step: 80, loss: 4.8918875108938664e-05
step: 90, loss: 4.056948819197714e-05
step: 100, loss: 2.8836408091592602e-05
step: 110, loss: 0.00010508413834031671
step: 120, loss: 2.10472153412411e-05
step: 130, loss: 3.130987170152366e-05
step: 140, loss: 5.025444625061937e-05
step: 150, loss: 3.061704046558589e-05
step: 160, loss: 5.123848677612841e-05
step: 170, loss: 1.8104710761690512e-05
step: 180, loss: 8.665960922371596e-05
step: 190, loss: 0.0002585176262073219
step: 200, loss: 2.323032276763115e-05
step: 210, loss: 4.545311094261706e-05
step: 220, loss: 0.00011735848238458857
step: 230, loss: 5.701044938177802e-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=0.9865168539325843, f1=0.9809203142536477,

```

```

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: 2.0432806195458397e-05
step: 10, loss: 1.6994519683066756e-05
step: 20, loss: 3.2982596167130396e-05
step: 30, loss: 3.873688183375634e-05
step: 40, loss: 9.175633022096008e-05
step: 50, loss: 0.0005801132647320628
step: 60, loss: 0.0007629996980540454
step: 70, loss: 6.5269174228888e-05
step: 80, loss: 9.87551175057888e-05
step: 90, loss: 9.640400094212964e-05
step: 100, loss: 3.672143429866992e-05
step: 110, loss: 2.6437730411998928e-05
step: 120, loss: 3.5504650440998375e-05
step: 130, loss: 5.5714754125801846e-05
step: 140, loss: 2.307392242073547e-05
step: 150, loss: 4.5490218326449394e-05
step: 160, loss: 2.3014310500002466e-05
step: 170, loss: 2.1330584786483087e-05
step: 180, loss: 4.425414226716384e-05
step: 190, loss: 0.00027412123745307326
step: 200, loss: 9.030352521222085e-05
step: 210, loss: 2.9872273444198072e-05
step: 220, loss: 6.0779890191042796e-05
step: 230, loss: 2.9138467652956024e-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=0.9842342342342343, f1=0.9820224719101124,
best_f1=0.9774266365688488
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:12, 194.56it/s]
load_f1 = 0.9876819708846584
real_f1 = 0.9865771812080537
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 238.20it/s]

```

2.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

step: 0, loss: 0.6267108917236328

step: 10, loss: 0.5949034690856934

step: 20, loss: 0.537598729133606

step: 30, loss: 0.20286037027835846
step: 40, loss: 0.1409284919500351
step: 50, loss: 0.19547297060489655
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.1010998785495758
step: 70, loss: 0.24856886267662048
step: 80, loss: 0.07708657532930374
step: 90, loss: 0.6445007920265198
step: 100, loss: 0.09472587704658508
step: 110, loss: 0.08808659017086029
step: 120, loss: 0.11816205084323883
step: 130, loss: 0.1672636717557907
step: 140, loss: 0.09022317826747894
step: 150, loss: 0.054991621524095535
step: 160, loss: 0.04236438125371933
step: 170, loss: 0.16752591729164124
step: 180, loss: 0.16008348762989044
step: 190, loss: 0.016779428347945213
step: 200, loss: 0.15533176064491272
step: 210, loss: 0.02326021157205105
step: 220, loss: 0.21525363624095917
step: 230, loss: 0.1786002665758133
step: 240, loss: 0.13484081625938416
step: 250, loss: 0.029026880860328674
step: 260, loss: 0.12169017642736435
step: 270, loss: 0.01198496762663126
step: 280, loss: 0.07245443016290665
step: 290, loss: 0.06600148975849152
step: 300, loss: 0.0652301013469696
step: 310, loss: 0.12004338949918747
step: 320, loss: 0.14482572674751282
step: 330, loss: 0.02927692048251629
step: 340, loss: 0.07065629214048386
step: 350, loss: 0.05797342211008072
step: 360, loss: 0.05707342550158501
step: 370, loss: 0.10673362761735916
step: 380, loss: 0.058016881346702576
step: 390, loss: 0.11114884167909622
step: 400, loss: 0.3040558397769928
step: 410, loss: 0.17043234407901764
step: 420, loss: 0.06233002990484238
step: 430, loss: 0.1818757802248001
step: 440, loss: 0.040666092187166214
step: 450, loss: 0.020280905067920685
step: 460, loss: 0.03424810245633125
step: 470, loss: 0.1947079598903656
step: 480, loss: 0.08259745687246323
step: 490, loss: 0.057367559522390366

```

step: 500, loss: 0.047298043966293335
step: 510, loss: 0.07227537781000137
step: 520, loss: 0.09187664836645126
step: 530, loss: 0.005787322297692299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9255419415645617, f1=0.9123951537744641,
best_f1=0.9123951537744641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08365704864263535
step: 10, loss: 0.06841093301773071
step: 20, loss: 0.0053204321302473545
step: 30, loss: 0.01620025746524334
step: 40, loss: 0.10672890394926071
step: 50, loss: 0.24341875314712524
step: 60, loss: 0.02036689594388008
step: 70, loss: 0.049937907606363297
step: 80, loss: 0.08762089908123016
step: 90, loss: 0.033537160605192184
step: 100, loss: 0.03521634265780449
step: 110, loss: 0.0237546656280756
step: 120, loss: 0.08913103491067886
step: 130, loss: 0.06665819138288498
step: 140, loss: 0.011206564493477345
step: 150, loss: 0.030281338840723038
step: 160, loss: 0.05323518067598343
step: 170, loss: 0.09199122339487076
step: 180, loss: 0.023362373933196068
step: 190, loss: 0.17232123017311096
step: 200, loss: 0.014629095792770386
step: 210, loss: 0.04935980215668678
step: 220, loss: 0.08513402938842773
step: 230, loss: 0.009006624110043049
step: 240, loss: 0.03511694073677063
step: 250, loss: 0.04595979303121567
step: 260, loss: 0.012115775607526302
step: 270, loss: 0.3277590572834015
step: 280, loss: 0.10294631123542786
step: 290, loss: 0.032370004802942276
step: 300, loss: 0.062254685908555984
step: 310, loss: 0.016098717227578163

```

```

step: 320, loss: 0.18581122159957886
step: 330, loss: 0.07280951738357544
step: 340, loss: 0.01679197885096073
step: 350, loss: 0.0034065269865095615
step: 360, loss: 0.16427353024482727
step: 370, loss: 0.16442745923995972
step: 380, loss: 0.07911697030067444
step: 390, loss: 0.06360714882612228
step: 400, loss: 0.13753831386566162
step: 410, loss: 0.05885427072644234
step: 420, loss: 0.028204575181007385
step: 430, loss: 0.012115572579205036
step: 440, loss: 0.14023113250732422
step: 450, loss: 0.01886637508869171
step: 460, loss: 0.030413448810577393
step: 470, loss: 0.13205039501190186
step: 480, loss: 0.27890586853027344
step: 490, loss: 0.03020332008600235
step: 500, loss: 0.14858897030353546
step: 510, loss: 0.023049695417284966
step: 520, loss: 0.06773140281438828
step: 530, loss: 0.03608666732907295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9322892676186089, f1=0.9237170596393898,
best_f1=0.9237170596393898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08359236270189285
step: 10, loss: 0.010207245126366615
step: 20, loss: 0.13499143719673157
step: 30, loss: 0.051113300025463104
step: 40, loss: 0.0034615250770002604
step: 50, loss: 0.07552739977836609
step: 60, loss: 0.011022055521607399
step: 70, loss: 0.004647609777748585
step: 80, loss: 0.0068703521974384785
step: 90, loss: 0.015369576402008533
step: 100, loss: 0.013424168340861797
step: 110, loss: 0.001357737579382956
step: 120, loss: 0.010861239396035671
step: 130, loss: 0.026779232546687126

```

```

step: 140, loss: 0.14863890409469604
step: 150, loss: 0.04681997746229172
step: 160, loss: 0.04432179778814316
step: 170, loss: 0.02940961718559265
step: 180, loss: 0.1163279190659523
step: 190, loss: 0.007248768582940102
step: 200, loss: 0.017974290996789932
step: 210, loss: 0.15389969944953918
step: 220, loss: 0.1935657262802124
step: 230, loss: 0.08883057534694672
step: 240, loss: 0.0022393178660422564
step: 250, loss: 0.05688856169581413
step: 260, loss: 0.024937931448221207
step: 270, loss: 0.015651509165763855
step: 280, loss: 0.14605003595352173
step: 290, loss: 0.04556634649634361
step: 300, loss: 0.024985909461975098
step: 310, loss: 0.02086411975324154
step: 320, loss: 0.04578321427106857
step: 330, loss: 0.004410558380186558
step: 340, loss: 0.012493566609919071
step: 350, loss: 0.06483016908168793
step: 360, loss: 0.03621957078576088
step: 370, loss: 0.015281269326806068
step: 380, loss: 0.022103238850831985
step: 390, loss: 0.14312724769115448
step: 400, loss: 0.05670670419931412
step: 410, loss: 0.023675231263041496
step: 420, loss: 0.07252579182386398
step: 430, loss: 0.0021967205684632063
step: 440, loss: 0.0011317782336845994
step: 450, loss: 0.0895102322101593
step: 460, loss: 0.06461960077285767
step: 470, loss: 0.037298426032066345
step: 480, loss: 0.005868798587471247
step: 490, loss: 0.01004727091640234
step: 500, loss: 0.02385173924267292
step: 510, loss: 0.004255097359418869
step: 520, loss: 0.0610102042555809
step: 530, loss: 0.022710241377353668
/usr/local/lib/python3.7/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=0.931098696461825, f1=0.9208566108007449,
best_f1=0.9237170596393898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```


step: 0, loss: 0.10373334586620331
step: 10, loss: 0.004538115579634905
step: 20, loss: 0.0034034624695777893
step: 30, loss: 0.0018793201306834817
step: 40, loss: 0.00446068961173296
step: 50, loss: 0.0040495614521205425
step: 60, loss: 0.0013260379200801253
step: 70, loss: 0.04634769633412361
step: 80, loss: 0.04119456559419632
step: 90, loss: 0.11648828536272049
step: 100, loss: 0.020329363644123077
step: 110, loss: 0.007819732651114464
step: 120, loss: 0.0007073513115756214
step: 130, loss: 0.0025236376095563173
step: 140, loss: 0.0013299854472279549
step: 150, loss: 0.002994928741827607
step: 160, loss: 0.02132030762732029
step: 170, loss: 0.009368862956762314
step: 180, loss: 0.05705799162387848
step: 190, loss: 0.10807376354932785
step: 200, loss: 0.026253771036863327
step: 210, loss: 0.06667780131101608
step: 220, loss: 0.052184831351041794
step: 230, loss: 0.11344876885414124
step: 240, loss: 0.009141925722360611
step: 250, loss: 0.014909694902598858
step: 260, loss: 0.004042335320264101
step: 270, loss: 0.12369564920663834
step: 280, loss: 0.24549490213394165
step: 290, loss: 0.06653625518083572
step: 300, loss: 0.00686072651296854
step: 310, loss: 0.017623314633965492
step: 320, loss: 0.004089657217264175
step: 330, loss: 0.02255670167505741
step: 340, loss: 0.015692180022597313
step: 350, loss: 0.012974699027836323
step: 360, loss: 0.04118644446134567
step: 370, loss: 0.001313969842158258
step: 380, loss: 0.05928583815693855
step: 390, loss: 0.0042972867377102375
step: 400, loss: 0.0004885734524577856
step: 410, loss: 0.015492245554924011
step: 420, loss: 0.001211880473420024
step: 430, loss: 0.07640738785266876
step: 440, loss: 0.01801920495927334
step: 450, loss: 0.007149477023631334
step: 460, loss: 0.11699376255273819
step: 470, loss: 0.0652945339679718

```

step: 480, loss: 0.026194222271442413
step: 490, loss: 0.002678891411051154
step: 500, loss: 0.12307673692703247
step: 510, loss: 0.025806767866015434
step: 520, loss: 0.1650703251361847
step: 530, loss: 0.032742295414209366
/usr/local/lib/python3.7/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=0.9289667896678966, f1=0.924074074074074,
best_f1=0.9237170596393898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02278817445039749
step: 10, loss: 0.026878682896494865
step: 20, loss: 0.003555754665285349
step: 30, loss: 0.0232352614402771
step: 40, loss: 0.0010125263361260295
step: 50, loss: 0.0012113350676372647
step: 60, loss: 0.024184565991163254
step: 70, loss: 0.002788554411381483
step: 80, loss: 0.0006838064291514456
step: 90, loss: 0.0021595419384539127
step: 100, loss: 0.001983568537980318
step: 110, loss: 0.0037637418136000633
step: 120, loss: 0.0010938530322164297
step: 130, loss: 0.0006932224496267736
step: 140, loss: 0.009436097927391529
step: 150, loss: 0.0005007949657738209
step: 160, loss: 0.0007670138729736209
step: 170, loss: 0.03792068362236023
step: 180, loss: 0.0010052559664472938
step: 190, loss: 0.05360248312354088
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.00571548892185092
step: 210, loss: 0.001392226666212082
step: 220, loss: 0.0020030708983540535
step: 230, loss: 0.006186855956912041
step: 240, loss: 0.007948512211441994
step: 250, loss: 0.0006830174825154245
step: 260, loss: 0.13553956151008606
step: 270, loss: 0.0014284623321145773
step: 280, loss: 0.018283482640981674
step: 290, loss: 0.1843915730714798
step: 300, loss: 0.00264290114864707
step: 310, loss: 0.029399879276752472
step: 320, loss: 0.030917946249246597

```

```

step: 330, loss: 0.07687973231077194
step: 340, loss: 0.037389595061540604
step: 350, loss: 0.05894426628947258
step: 360, loss: 0.03917621821165085
step: 370, loss: 0.00036205153446644545
step: 380, loss: 0.00011727273522410542
step: 390, loss: 0.0003319418174214661
step: 400, loss: 0.012582086957991123
step: 410, loss: 0.00028513517463579774
step: 420, loss: 0.0013155374908819795
step: 430, loss: 0.012906249612569809
step: 440, loss: 0.004981799051165581
step: 450, loss: 0.009254483506083488
step: 460, loss: 0.002707729348912835
step: 470, loss: 0.017354723066091537
step: 480, loss: 0.03092269040644169
step: 490, loss: 0.0003086405631620437
step: 500, loss: 0.0026247436180710793
step: 510, loss: 0.00996307097375393
step: 520, loss: 0.0031685843132436275
step: 530, loss: 0.00024300329096149653
/usr/local/lib/python3.7/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=0.9317972350230416, f1=0.925672594619243,
best_f1=0.9237170596393898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0065737827681005
step: 10, loss: 0.0005345235113054514
step: 20, loss: 0.10865435004234314
step: 30, loss: 0.004927552305161953
step: 40, loss: 0.008806237950921059
step: 50, loss: 0.02081051468849182
step: 60, loss: 0.0022196476347744465
step: 70, loss: 0.00016087025869637728
step: 80, loss: 0.037138037383556366
step: 90, loss: 0.0004853571008425206
step: 100, loss: 0.006361798848956823
step: 110, loss: 0.01635398343205452
step: 120, loss: 0.006796635687351227
step: 130, loss: 0.0002715991286095232
step: 140, loss: 0.013157746754586697
step: 150, loss: 0.0009171867277473211
step: 160, loss: 0.0012752545299008489
step: 170, loss: 0.015362939797341824
step: 180, loss: 0.0023020405787974596

```

```

step: 190, loss: 0.0016843498451635242
step: 200, loss: 0.00338633987121284
step: 210, loss: 0.1440388262271881
step: 220, loss: 0.003238001372665167
step: 230, loss: 0.0007070439751259983
step: 240, loss: 0.022936314344406128
step: 250, loss: 0.000445208657765761
step: 260, loss: 0.00014578555419575423
step: 270, loss: 0.052812717854976654
step: 280, loss: 0.0003088225203100592
step: 290, loss: 0.0002249312528874725
step: 300, loss: 0.0012863962911069393
step: 310, loss: 0.0015991046093404293
step: 320, loss: 0.0007756870472803712
step: 330, loss: 0.0006909077055752277
step: 340, loss: 0.008406628854572773
step: 350, loss: 0.00630797678604722
step: 360, loss: 0.008464646525681019
step: 370, loss: 0.04009654000401497
step: 380, loss: 0.00025772955268621445
step: 390, loss: 0.0069751739501953125
step: 400, loss: 0.0179547518491745
step: 410, loss: 0.0012363995192572474
step: 420, loss: 0.020266206935048103
step: 430, loss: 0.00047405078657902777
step: 440, loss: 0.001322952681221068
step: 450, loss: 0.010142105631530285
step: 460, loss: 0.0020453513134270906
step: 470, loss: 0.0002190140658058226
step: 480, loss: 0.16267129778862
step: 490, loss: 0.004592524841427803
step: 500, loss: 0.0008477346273139119
step: 510, loss: 0.0028941738419234753
step: 520, loss: 0.0016385759226977825
step: 530, loss: 0.013170318678021431
/usr/local/lib/python3.7/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=0.9283694627709709, f1=0.9169027384324836,
best_f1=0.9237170596393898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034794455859810114
step: 10, loss: 0.0016164719127118587
step: 20, loss: 0.01802416332066059
step: 30, loss: 0.0010625554714351892
step: 40, loss: 0.018606388941407204

```

step: 50, loss: 0.0017768593970686197
step: 60, loss: 0.0007774241385050118
step: 70, loss: 0.0033670098055154085
step: 80, loss: 0.0013316841796040535
step: 90, loss: 0.008130455389618874
step: 100, loss: 8.258473098976538e-05
step: 110, loss: 0.0006110292160883546
step: 120, loss: 0.0005417084903456271
step: 130, loss: 0.03722330555319786
step: 140, loss: 0.00021566047507803887
step: 150, loss: 0.0025025310460478067
step: 160, loss: 0.0028624781407415867
step: 170, loss: 0.00036270826240070164
step: 180, loss: 0.0002417118230368942
step: 190, loss: 0.009742464870214462
step: 200, loss: 0.0003952060069423169
step: 210, loss: 0.008752459660172462
step: 220, loss: 0.00048724241787567735
step: 230, loss: 0.02424110844731331
step: 240, loss: 0.0006347007001750171
step: 250, loss: 0.0011045702267438173
step: 260, loss: 0.0006574333528988063
step: 270, loss: 0.00018273526802659035
step: 280, loss: 0.0001252698857570067
step: 290, loss: 7.222445856314152e-05
step: 300, loss: 0.0034550102427601814
step: 310, loss: 0.00046179533819667995
step: 320, loss: 0.012091142125427723
step: 330, loss: 0.000801367626991123
step: 340, loss: 0.02671816013753414
step: 350, loss: 0.0021485567558556795
step: 360, loss: 0.002905531320720911
step: 370, loss: 0.0015133125707507133
step: 380, loss: 0.004425894469022751
step: 390, loss: 0.00013642046542372555
step: 400, loss: 0.004521467722952366
step: 410, loss: 0.0026388177648186684
step: 420, loss: 0.0006276018102653325
step: 430, loss: 0.0006445940234698355
step: 440, loss: 0.01788366213440895
step: 450, loss: 0.03736544027924538
step: 460, loss: 0.0007683420553803444
step: 470, loss: 0.00031377305276691914
step: 480, loss: 0.10048018395900726
step: 490, loss: 0.00034412628156132996
step: 500, loss: 0.015656357631087303
step: 510, loss: 0.002229053294286132
step: 520, loss: 0.00431671692058444

```

step: 530, loss: 0.023715835064649582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9333333333333335, f1=0.9186155285313378,
best_f1=0.9186155285313378
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003201083280146122
step: 10, loss: 0.013619175180792809
step: 20, loss: 3.9743554225424305e-05
step: 30, loss: 0.00831609033048153
step: 40, loss: 6.029223368386738e-05
step: 50, loss: 0.005945098586380482
step: 60, loss: 8.064838766586035e-05
step: 70, loss: 0.0001460518251406029
step: 80, loss: 6.690275768050924e-05
step: 90, loss: 0.002073865383863449
step: 100, loss: 0.0010674545774236321
step: 110, loss: 0.002372844610363245
step: 120, loss: 0.003513084724545479
step: 130, loss: 0.007938513532280922
step: 140, loss: 0.008363411761820316
step: 150, loss: 0.015068779699504375
step: 160, loss: 0.005752606317400932
step: 170, loss: 0.00035815523006021976
step: 180, loss: 0.00017098613898269832
step: 190, loss: 0.0008772542933002114
step: 200, loss: 0.0005622310563921928
step: 210, loss: 0.0004202675190754235
step: 220, loss: 0.0028241737745702267
step: 230, loss: 0.002016836777329445
step: 240, loss: 6.76106647006236e-05
step: 250, loss: 5.2067516662646085e-05
step: 260, loss: 0.0014074469218030572
step: 270, loss: 0.00014177229604683816
step: 280, loss: 0.001087796757929027
step: 290, loss: 4.7807174269109964e-05
step: 300, loss: 0.001212200615555048
step: 310, loss: 0.04895501211285591
step: 320, loss: 0.01335106696933508
step: 330, loss: 0.0006569130928255618
step: 340, loss: 0.045484092086553574

```

```

step: 350, loss: 5.994329330860637e-05
step: 360, loss: 0.0010468154214322567
step: 370, loss: 0.0018980716122314334
step: 380, loss: 0.001933484454639256
step: 390, loss: 0.1753392219543457
step: 400, loss: 0.0007920280331745744
step: 410, loss: 0.000422172830440104
step: 420, loss: 0.002870857482776046
step: 430, loss: 0.050905413925647736
step: 440, loss: 0.0009701474336907268
step: 450, loss: 0.001071331207640469
step: 460, loss: 0.00017632299568504095
step: 470, loss: 0.0001672228827374056
step: 480, loss: 0.00024410145124420524
step: 490, loss: 0.0022089937701821327
step: 500, loss: 0.00020491628674790263
step: 510, loss: 7.92268110672012e-05
step: 520, loss: 0.002719321520999074
step: 530, loss: 0.00013802951434627175
/usr/local/lib/python3.7/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=0.929245283018868, f1=0.9136622390891841,
best_f1=0.9186155285313378
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015009976923465729
step: 10, loss: 0.004459016490727663
step: 20, loss: 0.005109467543661594
step: 30, loss: 0.0014643603935837746
step: 40, loss: 0.00047006004024297
step: 50, loss: 0.0033161654137074947
step: 60, loss: 0.014511534944176674
step: 80, loss: 0.11693661659955978
step: 90, loss: 0.006935563869774342
step: 100, loss: 9.295057679992169e-05
step: 110, loss: 0.0016391624230891466
step: 120, loss: 9.841416613198817e-05
step: 130, loss: 0.00726278917863965
step: 140, loss: 0.002244033617898822
step: 150, loss: 0.00258798711001873
step: 160, loss: 0.0017994772642850876
step: 170, loss: 0.018214615061879158
step: 180, loss: 6.131717236712575e-05
step: 190, loss: 3.4411226806696504e-05
step: 200, loss: 4.07157713198103e-05
step: 210, loss: 2.6005658583017066e-05

```

```

step: 220, loss: 0.005027747247368097
step: 230, loss: 2.584553658380173e-05
step: 240, loss: 0.015478532761335373
step: 250, loss: 0.00032200239365920424
step: 260, loss: 0.02133927121758461
step: 270, loss: 0.002294197678565979
step: 280, loss: 0.03641210123896599
step: 290, loss: 0.09670640528202057
step: 300, loss: 4.515824548434466e-05
step: 310, loss: 0.0058558788150548935
step: 320, loss: 4.121856181882322e-05
step: 330, loss: 0.0007868953980505466
step: 340, loss: 0.0001785172789823264
step: 350, loss: 0.0008069216273725033
step: 360, loss: 2.866518116206862e-05
step: 370, loss: 0.0004883807268925011
step: 380, loss: 0.0003768344467971474
step: 390, loss: 0.001467496738769114
step: 400, loss: 0.00011901782272616401
step: 410, loss: 0.0002845374692697078
step: 420, loss: 3.6570054362528026e-05
step: 430, loss: 5.688723103958182e-05
step: 440, loss: 0.00017060598474927247
step: 450, loss: 0.013632423244416714
step: 460, loss: 0.0005729796830564737
step: 470, loss: 4.5702006900683045e-05
step: 480, loss: 0.0001354449923383072
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 490, loss: 0.01247299276292324
step: 500, loss: 0.01146673783659935
step: 510, loss: 0.01349383220076561
step: 520, loss: 0.0001955333718797192
step: 530, loss: 0.00011340740456944332
/usr/local/lib/python3.7/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=0.9284403669724771, f1=0.9206785878037598,
best_f1=0.9186155285313378
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036461438867263496
step: 10, loss: 0.00027924214373342693
step: 20, loss: 6.603400834137574e-05
step: 30, loss: 5.4604806791758165e-05
step: 40, loss: 0.0002597000857349485
step: 50, loss: 7.475588790839538e-05
step: 60, loss: 5.7134351663989946e-05

```


step: 70, loss: 0.00045530160423368216
step: 80, loss: 0.00036486340104602277
step: 90, loss: 0.00010051800200017169
step: 100, loss: 4.534426625468768e-05
step: 110, loss: 4.4604792492464185e-05
step: 120, loss: 5.922196578467265e-05
step: 130, loss: 0.0035334366839379072
step: 140, loss: 0.007226415444165468
step: 150, loss: 3.420051507418975e-05
step: 160, loss: 4.6549906983273104e-05
step: 170, loss: 8.128438639687374e-05
step: 180, loss: 0.0001973129838006571
step: 190, loss: 0.00019776643603108823
step: 200, loss: 0.017376836389303207
step: 210, loss: 0.00016071309801191092
step: 220, loss: 0.0024476235266774893
step: 230, loss: 0.0029480597004294395
step: 240, loss: 0.00037060491740703583
step: 250, loss: 4.026342139695771e-05
step: 260, loss: 0.0007166177383624017
step: 270, loss: 0.00011378271301509812
step: 280, loss: 0.00022544633247889578
step: 290, loss: 0.0010113168973475695
step: 300, loss: 5.201736348681152e-05
step: 310, loss: 0.00014898173685651273
step: 320, loss: 0.003309030784294009
step: 330, loss: 0.015303228050470352
step: 340, loss: 5.912263804930262e-05
step: 350, loss: 0.00032994477078318596
step: 360, loss: 0.00015864797751419246
step: 370, loss: 0.0009129879763349891
step: 380, loss: 0.01794787123799324
step: 390, loss: 0.0008043509442359209
step: 400, loss: 7.359973096754402e-05
step: 410, loss: 0.00040358869591727853
step: 420, loss: 7.559070218121633e-05
step: 430, loss: 2.4753557227086276e-05
step: 440, loss: 0.0003825908061116934
step: 450, loss: 2.291007331223227e-05
step: 460, loss: 4.147109211771749e-05
step: 470, loss: 0.0018764908891171217
step: 480, loss: 9.932081593433395e-05
step: 490, loss: 0.001043273019604385
step: 500, loss: 0.00032499467488378286
step: 510, loss: 6.480012234533206e-05
step: 520, loss: 0.00010890231351368129
step: 530, loss: 0.000976848998107016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9335810496980957, f1=0.9210770659238626,
best_f1=0.9210770659238626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004673333023674786
step: 10, loss: 0.00018405658192932606
step: 20, loss: 0.00027349789161235094
step: 30, loss: 0.0022995942272245884
step: 40, loss: 4.417878881213255e-05
step: 50, loss: 7.078459020704031e-05
step: 60, loss: 0.030444126576185226
step: 70, loss: 6.07430592935998e-05
step: 80, loss: 7.179356180131435e-05
step: 90, loss: 2.980122917506378e-05
step: 100, loss: 9.953399421647191e-05
step: 110, loss: 0.0018786047585308552
step: 120, loss: 0.0007848816458135843
step: 130, loss: 2.686226252990309e-05
step: 140, loss: 5.303096259012818e-05
step: 150, loss: 2.961434620374348e-05
step: 160, loss: 2.258963286294602e-05
step: 170, loss: 4.469109990168363e-05
step: 180, loss: 0.00013972094166092575
step: 190, loss: 0.00011699800961650908
step: 200, loss: 0.0035568636376410723
step: 210, loss: 0.00041207001777365804
step: 220, loss: 0.00029812383581884205
step: 230, loss: 2.7895292078028433e-05
step: 240, loss: 0.00010944154200842604
step: 250, loss: 1.9803344912361354e-05
step: 260, loss: 3.990680488641374e-05
step: 270, loss: 1.9043418433284387e-05
step: 280, loss: 2.0354636944830418e-05
step: 290, loss: 0.00010017187014454976
step: 300, loss: 2.136416151188314e-05
step: 310, loss: 0.00013007030065637082
step: 320, loss: 0.0001284262107219547
step: 330, loss: 4.360602542874403e-05
step: 340, loss: 2.3971755581442267e-05
step: 350, loss: 2.2112837541499175e-05
step: 360, loss: 1.724416870274581e-05

```

```
step: 370, loss: 0.00012031526421196759
step: 380, loss: 2.0168414266663603e-05
step: 390, loss: 1.8257402189192362e-05
step: 400, loss: 7.956608169479296e-05
step: 410, loss: 2.5212009859387763e-05
step: 420, loss: 5.0864924560301006e-05
step: 430, loss: 2.4455908715026453e-05
step: 440, loss: 0.0015506744384765625
step: 450, loss: 0.00018661464855540544
step: 460, loss: 0.024676527827978134
step: 470, loss: 0.0002514492080081254
step: 480, loss: 6.737290823366493e-05
step: 490, loss: 0.0016579435905441642
step: 500, loss: 2.8847591238445602e-05
step: 510, loss: 0.00016154276090674102
step: 520, loss: 2.7115675038658082e-05
step: 530, loss: 0.006052315700799227
/usr/local/lib/python3.7/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=0.9128986197049024, f1=0.9021276595744682,
best_f1=0.9210770659238626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.091812686761841e-05
step: 10, loss: 2.7230957130086608e-05
step: 20, loss: 0.0003689429140649736
step: 30, loss: 0.00038914658944122493
step: 40, loss: 4.2253010178683326e-05
step: 50, loss: 0.010094981640577316
step: 60, loss: 0.00028922504861839116
step: 70, loss: 0.09121192991733551
step: 80, loss: 4.472156797419302e-05
step: 90, loss: 3.781131454161368e-05
step: 100, loss: 0.0001077260312740691
step: 110, loss: 4.892785000265576e-05
step: 120, loss: 2.014217898249626e-05
step: 130, loss: 0.0001078677159966901
step: 140, loss: 1.8264876416651532e-05
step: 150, loss: 0.00020902951655443758
step: 160, loss: 2.245921132271178e-05
step: 170, loss: 2.1889327399549074e-05
step: 180, loss: 1.7262762412428856e-05
step: 190, loss: 3.1449566449737176e-05
step: 200, loss: 3.328000457258895e-05
step: 210, loss: 9.046077320817858e-05
step: 220, loss: 0.00024745232076384127
```

```

step: 230, loss: 1.4398111488844734e-05
step: 240, loss: 1.842115852923598e-05
step: 250, loss: 1.7426684280508198e-05
step: 260, loss: 1.8246219042339362e-05
step: 270, loss: 3.0452221835730597e-05
step: 280, loss: 0.0016516970936208963
step: 290, loss: 5.35022554686293e-05
step: 300, loss: 0.00013033865252509713
step: 310, loss: 4.2141480662394315e-05
step: 320, loss: 0.0007359344162978232
step: 330, loss: 0.0001113471225835383
step: 340, loss: 2.434396265016403e-05
step: 350, loss: 2.5279008696088567e-05
step: 360, loss: 7.45789147913456e-05
step: 370, loss: 0.00010545044642640278
step: 380, loss: 8.879639790393412e-05
step: 390, loss: 3.0470928322756663e-05
step: 400, loss: 2.6843141313293017e-05
step: 410, loss: 6.230978760868311e-05
step: 420, loss: 0.008535871282219887
step: 430, loss: 0.004883110523223877
step: 440, loss: 8.196626731660217e-05
step: 450, loss: 1.957236418093089e-05
step: 460, loss: 2.217959627159871e-05
step: 470, loss: 2.392696842434816e-05
step: 480, loss: 0.0001747265923768282
step: 490, loss: 5.8638572227209806e-05
step: 500, loss: 8.128750778269023e-05
step: 510, loss: 0.010971101000905037
step: 520, loss: 2.0283880076021887e-05
step: 530, loss: 1.8950269804918207e-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=0.9319129226493746, f1=0.9155514536225196,
best_f1=0.9210770659238626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013817923609167337
step: 10, loss: 2.231341932201758e-05
step: 20, loss: 2.0965158910257742e-05
step: 30, loss: 0.00023222551681101322
step: 40, loss: 2.5990402718889527e-05
step: 50, loss: 0.002507340395823121
step: 60, loss: 6.046736962161958e-05
step: 70, loss: 1.5005200111772865e-05
step: 80, loss: 0.00019808470096904784

```

```
step: 90, loss: 0.009199694730341434
step: 100, loss: 0.00030818433151580393
step: 110, loss: 3.0794595659244806e-05
step: 120, loss: 6.707698776153848e-05
step: 130, loss: 6.692158785881475e-05
step: 140, loss: 2.4291650333907455e-05
step: 150, loss: 9.939981828210875e-05
step: 160, loss: 3.8697737181792036e-05
step: 170, loss: 0.0002070335322059691
step: 180, loss: 2.6437093765707687e-05
step: 190, loss: 2.2790723960497417e-05
step: 200, loss: 0.00023740452888887376
step: 210, loss: 6.475460395449772e-05
step: 220, loss: 2.037317244685255e-05
step: 230, loss: 1.545974191685673e-05
step: 240, loss: 0.00011184228060301393
step: 250, loss: 2.5900690161506645e-05
step: 260, loss: 0.0006422192673198879
step: 270, loss: 1.5646022802684456e-05
step: 280, loss: 2.01050061150454e-05
step: 290, loss: 4.283800080884248e-05
step: 300, loss: 4.5759774366160855e-05
step: 310, loss: 0.00046916812425479293
step: 320, loss: 0.012161369435489178
step: 330, loss: 9.897667041514069e-05
step: 340, loss: 2.699204742384609e-05
step: 350, loss: 2.055190270766616e-05
step: 360, loss: 6.827714969404042e-05
step: 370, loss: 0.00039745672256685793
step: 380, loss: 3.3062828151741996e-05
step: 390, loss: 0.0006105850916355848
step: 400, loss: 1.4833929526503198e-05
step: 410, loss: 5.8136738516623154e-05
step: 420, loss: 3.1356499675894156e-05
step: 430, loss: 0.025950388982892036
step: 440, loss: 0.00046087041846476495
step: 450, loss: 4.4768996303901076e-05
step: 460, loss: 0.00022907556558493525
step: 470, loss: 1.3023494830122218e-05
step: 480, loss: 0.0021996928844600916
step: 490, loss: 1.4953160643926822e-05
step: 500, loss: 1.4375740647665225e-05
step: 510, loss: 0.000272689969278872
step: 520, loss: 5.897458322579041e-05
step: 530, loss: 1.45769226946868e-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=0.9326568265682658, f1=0.9272559852670351,  
best_f1=0.9210770659238626  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0004720010911114514  
step: 10, loss: 1.6942136426223442e-05  
step: 20, loss: 1.5057472410262562e-05  
step: 30, loss: 1.5996216461644508e-05  
step: 40, loss: 3.155082595185377e-05  
step: 50, loss: 5.1421789976302534e-05  
step: 60, loss: 1.534052898932714e-05  
step: 70, loss: 5.351890285965055e-05  
step: 80, loss: 1.9516402971930802e-05  
step: 90, loss: 1.6037187378969975e-05  
step: 100, loss: 5.179142317501828e-05  
step: 110, loss: 8.465362043352798e-05  
step: 120, loss: 0.0007849299581721425  
step: 130, loss: 0.0007349152583628893  
step: 140, loss: 1.3999421753396746e-05  
step: 150, loss: 1.424533638783032e-05  
step: 160, loss: 3.7858208088437095e-05  
step: 170, loss: 2.01049551833421e-05  
step: 180, loss: 2.7885762392543256e-05  
step: 190, loss: 3.897710848832503e-05  
step: 200, loss: 2.6024094040622003e-05  
step: 210, loss: 2.0622615920729004e-05  
step: 220, loss: 2.1162755729164928e-05  
step: 230, loss: 0.0002530695637688041  
step: 240, loss: 1.793688534235116e-05  
step: 250, loss: 3.113370985374786e-05  
step: 260, loss: 1.7150934581877664e-05  
step: 270, loss: 1.8223745428258553e-05  
step: 280, loss: 1.5429965060320683e-05  
step: 290, loss: 0.002830990357324481  
step: 300, loss: 2.3163302103057504e-05  
step: 310, loss: 2.759905328275636e-05  
step: 320, loss: 0.00013944787497166544  
step: 330, loss: 0.00021872615616302937  
step: 340, loss: 2.135276736225933e-05  
step: 350, loss: 0.00016185578715521842  
step: 360, loss: 8.107214671326801e-05  
step: 370, loss: 1.632774728932418e-05  
step: 380, loss: 1.884586708911229e-05  
step: 390, loss: 0.01989993080496788  
step: 400, loss: 6.906870839884505e-05  
step: 410, loss: 0.00012334735947661102  
step: 420, loss: 1.4945592738513369e-05
```

```
step: 430, loss: 0.00012879847781732678
step: 440, loss: 3.296313661849126e-05
step: 450, loss: 3.662180824903771e-05
step: 460, loss: 4.581890971167013e-05
step: 470, loss: 1.584719393576961e-05
step: 480, loss: 1.540762241347693e-05
step: 490, loss: 1.3805797607346904e-05
step: 500, loss: 2.1333824406610802e-05
step: 510, loss: 5.187015995034017e-05
step: 520, loss: 2.867431794584263e-05
step: 530, loss: 1.7042975741787814e-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=0.9330254041570438, f1=0.9259088817303267,
best_f1=0.9210770659238626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.1758929506177083e-05
step: 10, loss: 2.9170572815928608e-05
step: 20, loss: 7.11199245415628e-05
step: 30, loss: 3.376559106982313e-05
step: 40, loss: 0.03275738283991814
step: 50, loss: 0.00016879431495908648
step: 60, loss: 1.573912231833674e-05
step: 70, loss: 6.833308725617826e-05
step: 80, loss: 1.8939061192213558e-05
step: 90, loss: 1.9550045180949382e-05
step: 100, loss: 2.9442182494676672e-05
step: 110, loss: 0.0009706204291433096
step: 120, loss: 1.8372795238974504e-05
step: 130, loss: 0.01536665577441454
step: 140, loss: 1.5090991837496404e-05
step: 150, loss: 2.212769140896853e-05
step: 160, loss: 1.80673068825854e-05
step: 170, loss: 1.6923711882554926e-05
step: 180, loss: 1.3217189916758798e-05
step: 190, loss: 1.875656562333461e-05
step: 200, loss: 5.893664274481125e-05
step: 210, loss: 4.6101344196358696e-05
step: 220, loss: 1.3232122000772506e-05
step: 230, loss: 1.6163761756615713e-05
step: 240, loss: 1.5552890545222908e-05
step: 250, loss: 0.0001290549844270572
step: 260, loss: 1.4833964996796567e-05
step: 270, loss: 2.9188067856011912e-05
step: 280, loss: 5.079788752482273e-05
```

```

step: 290, loss: 1.2203956430312246e-05
step: 300, loss: 1.6789625078672543e-05
step: 310, loss: 1.8108339645550586e-05
step: 320, loss: 1.1429090591263957e-05
step: 330, loss: 0.0001567424915265292
step: 340, loss: 1.2084746231266763e-05
step: 350, loss: 4.341976091382094e-05
step: 360, loss: 0.00012074995902366936
step: 370, loss: 1.304586749029113e-05
step: 380, loss: 1.8685715986066498e-05
step: 390, loss: 1.7542135537951253e-05
step: 400, loss: 1.5564097338938154e-05
step: 410, loss: 2.382571983616799e-05
step: 420, loss: 2.529362973291427e-05
step: 430, loss: 1.9862996850861236e-05
step: 440, loss: 2.0585421225405298e-05
step: 450, loss: 1.343327221547952e-05
step: 460, loss: 1.553054607938975e-05
step: 470, loss: 0.011828272603452206
step: 480, loss: 1.2781352779711597e-05
step: 490, loss: 2.0588862753356807e-05
step: 500, loss: 0.0005423870170488954
step: 510, loss: 7.42389092920348e-05
step: 520, loss: 1.3243295143183786e-05
step: 530, loss: 0.00010760275472421199
/usr/local/lib/python3.7/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=0.9315960912052118, f1=0.9213691026827011,
best_f1=0.9210770659238626
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 235.99it/s]
load_f1 = 0.9326568265682658
real_f1 = 0.9325946445060019
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 247.91it/s]

```

2.1.4 Amazon-Google - Running the matcher

```

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

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

An input tensor was not cuda.

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

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

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

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

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

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

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

```
step: 0, loss: 0.5440030097961426  
step: 10, loss: 0.3649098873138428  
step: 20, loss: 0.37240099906921387  
step: 30, loss: 0.33230623602867126  
step: 40, loss: 0.18460886180400848  
step: 50, loss: 0.4423784613609314  
step: 60, loss: 0.32871192693710327
```

```

step: 70, loss: 0.2285580337047577
step: 80, loss: 0.2080768346786499
step: 90, loss: 0.4342476725578308
step: 100, loss: 0.47407668828964233
step: 110, loss: 0.2829470634460449
step: 120, loss: 0.21110309660434723
step: 130, loss: 0.23878689110279083
step: 140, loss: 0.29347696900367737
step: 150, loss: 0.317021906375885
step: 160, loss: 0.36376747488975525
step: 170, loss: 0.29025399684906006
step: 180, loss: 0.12273057550191879
step: 190, loss: 0.193452849984169
step: 200, loss: 0.3508928716182709
step: 210, loss: 0.23597677052021027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5614035087719298, f1=0.5689655172413792,
best_f1=0.5689655172413792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12699858844280243
step: 10, loss: 0.23318365216255188
step: 20, loss: 0.20210103690624237
step: 30, loss: 0.07564438879489899
step: 40, loss: 0.19424326717853546
step: 50, loss: 0.15821956098079681
step: 60, loss: 0.37144994735717773
step: 70, loss: 0.08273259550333023
step: 80, loss: 0.19825883209705353
step: 90, loss: 0.10404521971940994
step: 100, loss: 0.03958449512720108
step: 110, loss: 0.1006186306476593
step: 120, loss: 0.14164909720420837
step: 130, loss: 0.07383047044277191
step: 140, loss: 0.1710171103477478
step: 150, loss: 0.20303505659103394
step: 160, loss: 0.1815839558839798
step: 170, loss: 0.19226394593715668
step: 180, loss: 0.24092772603034973
step: 190, loss: 0.2661977708339691
step: 200, loss: 0.0842076912522316

```

```

step: 210, loss: 0.11858850717544556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5917159763313609, f1=0.5462753950338601,
best_f1=0.5462753950338601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04840746149420738
step: 10, loss: 0.1365516483783722
step: 20, loss: 0.13600566983222961
step: 30, loss: 0.19504602253437042
step: 40, loss: 0.08289333432912827
step: 50, loss: 0.14241282641887665
step: 60, loss: 0.2181982696056366
step: 70, loss: 0.09731274843215942
step: 80, loss: 0.10290251672267914
step: 90, loss: 0.14192579686641693
step: 100, loss: 0.33874034881591797
step: 110, loss: 0.18699157238006592
step: 120, loss: 0.11052210628986359
step: 130, loss: 0.1144929826259613
step: 140, loss: 0.06761745363473892
step: 150, loss: 0.08655218034982681
step: 160, loss: 0.018297262489795685
step: 170, loss: 0.07961829006671906
step: 180, loss: 0.08599044382572174
step: 190, loss: 0.2849714159965515
step: 200, loss: 0.06346059590578079
step: 210, loss: 0.07434717565774918
/usr/local/lib/python3.7/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=0.5678571428571428, f1=0.5911708253358925,
best_f1=0.5462753950338601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18475817143917084
step: 10, loss: 0.023033276200294495
step: 20, loss: 0.04112129658460617
step: 30, loss: 0.09433787316083908
step: 40, loss: 0.018937567248940468

```

```

step: 50, loss: 0.09494917094707489
step: 60, loss: 0.10174626857042313
step: 70, loss: 0.12905968725681305
step: 80, loss: 0.10980512946844101
step: 90, loss: 0.021245649084448814
step: 100, loss: 0.22099100053310394
step: 110, loss: 0.09047479927539825
step: 120, loss: 0.09676244109869003
step: 130, loss: 0.04829426109790802
step: 140, loss: 0.28177881240844727
step: 150, loss: 0.16927450895309448
step: 160, loss: 0.06610520929098129
step: 170, loss: 0.07565861940383911
step: 180, loss: 0.27562159299850464
step: 190, loss: 0.07170344144105911
step: 200, loss: 0.06688819080591202
step: 210, loss: 0.10500502586364746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6055045871559633, f1=0.6117647058823529,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12514781951904297
step: 10, loss: 0.06471233814954758
step: 20, loss: 0.060600124299526215
step: 30, loss: 0.02451370097696781
step: 40, loss: 0.05610426515340805
step: 50, loss: 0.05432124435901642
step: 60, loss: 0.09194162487983704
step: 70, loss: 0.032662149518728256
step: 80, loss: 0.014814816415309906
step: 90, loss: 0.10023006051778793
step: 100, loss: 0.01369871012866497
step: 110, loss: 0.1510085165500641
step: 120, loss: 0.24709127843379974
step: 130, loss: 0.0726146399974823
step: 140, loss: 0.04704253748059273
step: 150, loss: 0.010099618695676327
step: 160, loss: 0.06307049095630646
step: 170, loss: 0.09620897471904755
step: 180, loss: 0.07120735943317413

```

```

step: 190, loss: 0.02753276377916336
step: 200, loss: 0.09862633794546127
step: 210, loss: 0.026248600333929062
/usr/local/lib/python3.7/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=0.5724508050089445, f1=0.5773584905660376,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025474242866039276
step: 10, loss: 0.02129279263317585
step: 20, loss: 0.017426205798983574
step: 30, loss: 0.004559664987027645
step: 40, loss: 0.01038987748324871
step: 50, loss: 0.013603251427412033
step: 60, loss: 0.08126476407051086
step: 70, loss: 0.01222559716552496
step: 80, loss: 0.04516410827636719
step: 90, loss: 0.08184677362442017
step: 100, loss: 0.007114486303180456
step: 110, loss: 0.01077716238796711
step: 120, loss: 0.0779375210404396
step: 130, loss: 0.038559380918741226
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.09721089154481888
step: 150, loss: 0.05030950903892517
step: 160, loss: 0.007983379065990448
step: 170, loss: 0.18982046842575073
step: 180, loss: 0.004162045661360025
step: 190, loss: 0.141052708029747
step: 200, loss: 0.01039651408791542
step: 210, loss: 0.0145646333694458
/usr/local/lib/python3.7/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=0.5733558178752107, f1=0.6014760147601476,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026243319734930992
step: 10, loss: 0.0070615350268781185
step: 20, loss: 0.002327214926481247
step: 30, loss: 0.004089314490556717
step: 40, loss: 0.0074804495088756084
step: 50, loss: 0.063073068857193

```

```

step: 60, loss: 0.006509866565465927
step: 70, loss: 0.0026540763210505247
step: 80, loss: 0.06291814893484116
step: 90, loss: 0.13094118237495422
step: 100, loss: 0.006774385925382376
step: 110, loss: 0.013750174082815647
step: 120, loss: 0.07354842871427536
step: 130, loss: 0.026994558051228523
step: 140, loss: 0.0009045247570611537
step: 150, loss: 0.02206403575837612
step: 160, loss: 0.04165973886847496
step: 170, loss: 0.010068725794553757
step: 180, loss: 0.010877634398639202
step: 190, loss: 0.054673872888088226
step: 200, loss: 0.024627571925520897
step: 210, loss: 0.027714211493730545
/usr/local/lib/python3.7/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=0.5505376344086022, f1=0.5766590389016019,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023481929674744606
step: 10, loss: 0.05235607922077179
step: 20, loss: 0.008814997971057892
step: 30, loss: 0.10136166214942932
step: 40, loss: 0.008834786713123322
step: 50, loss: 0.004313493147492409
step: 60, loss: 0.1309971958398819
step: 70, loss: 0.04848785325884819
step: 80, loss: 0.05818169564008713
step: 90, loss: 0.032669223845005035
step: 100, loss: 0.06089962273836136
step: 110, loss: 0.15886102616786957
step: 120, loss: 0.000463584583485499
step: 130, loss: 0.0009499543230049312
step: 140, loss: 0.019307415932416916
step: 150, loss: 0.011429060250520706
step: 160, loss: 0.10042153298854828
step: 170, loss: 0.048660844564437866
step: 180, loss: 0.02963724173605442
step: 190, loss: 0.012204153463244438
step: 200, loss: 0.013967757113277912
step: 210, loss: 0.17532303929328918
/usr/local/lib/python3.7/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=0.5505376344086022, f1=0.5648148148148148,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00126696249935776
step: 10, loss: 0.005847221240401268
step: 20, loss: 0.002153032924979925
step: 30, loss: 0.009359597228467464
step: 40, loss: 0.004303536377847195
step: 50, loss: 0.0038986795116215944
step: 60, loss: 0.0011633114190772176
step: 70, loss: 0.0003012100060004741
step: 80, loss: 0.006528536323457956
step: 90, loss: 0.00024999325978569686
step: 100, loss: 0.0007697666878812015
step: 110, loss: 0.0007009035907685757
step: 120, loss: 0.0023588063195347786
step: 130, loss: 0.0503397136926651
step: 140, loss: 0.03571253642439842
step: 150, loss: 0.025977741926908493
step: 160, loss: 0.0002715529699344188
step: 170, loss: 0.026951676234602928
step: 180, loss: 0.0010127744171768427
step: 190, loss: 0.00023233925458043814
step: 200, loss: 0.08402568846940994
step: 210, loss: 0.00030590081587433815
/usr/local/lib/python3.7/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=0.5475113122171946, f1=0.5183374083129585,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001829241868108511
step: 10, loss: 0.03247559443116188
step: 20, loss: 0.0034548365511000156
step: 30, loss: 0.026379264891147614
step: 40, loss: 0.0019312002696096897
step: 50, loss: 0.0023093463387340307
step: 60, loss: 0.007936541922390461
step: 70, loss: 0.027882449328899384
step: 80, loss: 0.020934203639626503
step: 90, loss: 0.018798677250742912
step: 100, loss: 0.001285456819459796
step: 110, loss: 0.00099326076451689

```



```

step: 120, loss: 0.00040824501775205135
step: 130, loss: 0.06399556249380112
step: 140, loss: 0.006179390475153923
step: 150, loss: 0.05373639613389969
step: 160, loss: 0.011188933625817299
step: 170, loss: 0.00047093594912439585
step: 180, loss: 0.01826479844748974
step: 190, loss: 0.007365668658167124
step: 200, loss: 0.0003004566242452711
step: 210, loss: 0.0018026158213615417
/usr/local/lib/python3.7/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=0.5514223194748359, f1=0.5200945626477541,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08688677102327347
step: 10, loss: 0.0032816154416650534
step: 20, loss: 0.062116414308547974
step: 30, loss: 0.021217109635472298
step: 40, loss: 0.00992058590054512
step: 50, loss: 0.00035569435567595065
step: 60, loss: 0.05002252757549286
step: 70, loss: 0.00019459602481219918
step: 80, loss: 0.02245876006782055
step: 90, loss: 0.06906408816576004
step: 100, loss: 0.05124524608254433
step: 110, loss: 0.05557585135102272
step: 120, loss: 0.10983394831418991
step: 130, loss: 0.009980249218642712
step: 140, loss: 0.0017171531217172742
step: 150, loss: 0.0002914982324000448
step: 160, loss: 0.0016110822325572371
step: 170, loss: 0.0007894173613749444
step: 180, loss: 0.008809255436062813
step: 190, loss: 0.0003635731991380453
step: 200, loss: 0.0005215155542828143
step: 210, loss: 0.00021260109497234225
/usr/local/lib/python3.7/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=0.556237218813906, f1=0.5410199556541021,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.008021493442356586
step: 10, loss: 0.0003806862805504352
step: 20, loss: 0.009002079255878925
step: 30, loss: 0.023637965321540833
step: 40, loss: 0.014173829928040504
step: 50, loss: 0.00023624954337719828
step: 60, loss: 0.0008529247716069221
step: 70, loss: 0.0009915397968143225
step: 80, loss: 0.00019916104793082923
step: 90, loss: 0.0009173038997687399
step: 100, loss: 0.0007308708154596388
step: 110, loss: 0.0011588402558118105
step: 120, loss: 0.00020391591533552855
step: 130, loss: 5.570537905441597e-05
step: 140, loss: 0.0001384284405503422
step: 150, loss: 0.0017212569946423173
step: 160, loss: 0.0009598872275091708
step: 170, loss: 0.027891570702195168
step: 180, loss: 0.0006142371566966176
step: 190, loss: 0.0005457022343762219
step: 200, loss: 0.0022309748455882072
step: 210, loss: 0.0009063105098903179
/usr/local/lib/python3.7/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=0.5537757437070938, f1=0.5108433734939758,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004439695330802351
step: 10, loss: 0.00022233730123843998
step: 20, loss: 0.000212273167562671
step: 30, loss: 0.003942206036299467
step: 40, loss: 0.00022550227004103363
step: 50, loss: 0.0011720783077180386
step: 60, loss: 0.0001657952379900962
step: 70, loss: 0.005045795347541571
step: 80, loss: 0.0006223821546882391
step: 90, loss: 8.864616393111646e-05
step: 100, loss: 0.0005395510233938694
step: 110, loss: 0.0004158349765930325
step: 120, loss: 0.013831651769578457
step: 130, loss: 0.0011209760559722781
step: 140, loss: 0.020069291815161705
step: 150, loss: 0.00033714628079906106
step: 160, loss: 0.002538904082030058
step: 170, loss: 0.00021410835324786603

```

```

step: 180, loss: 0.013256402686238289
step: 190, loss: 0.00013344272156246006
step: 200, loss: 0.0004320737789385021
step: 210, loss: 0.000921803992241621
/usr/local/lib/python3.7/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=0.5454545454545454, f1=0.5073891625615763,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.109771493356675e-05
step: 10, loss: 0.01786985620856285
step: 20, loss: 0.0002677197626326233
step: 30, loss: 0.00010261988791171461
step: 40, loss: 0.00018457029364071786
step: 50, loss: 0.003202210646122694
step: 60, loss: 0.00013534966274164617
step: 70, loss: 0.00021610375551972538
step: 80, loss: 9.151885751634836e-05
step: 90, loss: 0.00022632947366219014
step: 100, loss: 0.0004058340273331851
step: 110, loss: 0.0002646278589963913
step: 120, loss: 0.0012614403385668993
step: 130, loss: 0.0013197524240240455
step: 140, loss: 0.002712210873141885
step: 150, loss: 0.00015359441749751568
step: 160, loss: 8.029102173168212e-05
step: 170, loss: 0.00013165593554731458
step: 180, loss: 0.00010918774933088571
step: 190, loss: 0.00017607415793463588
step: 200, loss: 0.00012459333811420947
step: 210, loss: 0.00021588038362096995
/usr/local/lib/python3.7/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=0.5433255269320844, f1=0.49131513647642683,
best_f1=0.6117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001342910254606977
step: 10, loss: 0.0001448486582376063
step: 20, loss: 0.00037971141864545643
step: 30, loss: 0.005125465337187052
step: 40, loss: 0.0001745152403600514
step: 50, loss: 0.00019600069208536297

```

```

step: 60, loss: 0.00015912462549749762
step: 70, loss: 0.00023602027795277536
step: 80, loss: 0.00044595947838388383
step: 90, loss: 0.00021244485105853528
step: 100, loss: 0.15458223223686218
step: 110, loss: 8.595974941272289e-05
step: 120, loss: 8.333199366461486e-05
step: 130, loss: 0.0007872087880969048
step: 140, loss: 8.104727021418512e-05
step: 150, loss: 0.0001492879819124937
step: 160, loss: 9.321113611804321e-05
step: 170, loss: 0.00012795823568012565
step: 180, loss: 9.38310768106021e-05
step: 190, loss: 0.0023372452706098557
step: 200, loss: 0.0009252224117517471
step: 210, loss: 0.0002357100456720218
/usr/local/lib/python3.7/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=0.5404157043879908, f1=0.5, best_f1=0.6117647058823529
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, 406.06it/s]
load_f1 = 0.6000000000000001
real_f1 = 0.6030534351145038
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 243.20it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

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

Defaults for this optimization level are:

```

enabled          : True
opt_level        : O2

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 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: 0.5532392859458923
step: 10, loss: 0.34067070484161377
step: 20, loss: 0.2793228328227997
step: 30, loss: 0.44483980536460876
step: 40, loss: 0.4168815612792969
step: 50, loss: 0.3100985288619995
step: 60, loss: 0.25709015130996704
step: 70, loss: 0.2888796031475067
step: 80, loss: 0.385023832321167
step: 90, loss: 0.28193479776382446
step: 100, loss: 0.3175031244754791
step: 110, loss: 0.3635217249393463
step: 120, loss: 0.13344869017601013
step: 130, loss: 0.1476389765739441
step: 140, loss: 0.09008416533470154
step: 150, loss: 0.1557929664850235
step: 160, loss: 0.12676629424095154
step: 170, loss: 0.2259700745344162

```

```

step: 180, loss: 0.03144976124167442
step: 190, loss: 0.18644851446151733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6237623762376238, f1=0.6666666666666667,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2478090524673462
step: 10, loss: 0.15965536236763
step: 20, loss: 0.1556958109140396
step: 30, loss: 0.15573255717754364
step: 40, loss: 0.07999138534069061
step: 50, loss: 0.03592995926737785
step: 60, loss: 0.26432088017463684
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.20749181509017944
step: 80, loss: 0.14840617775917053
step: 90, loss: 0.173637256026268
step: 100, loss: 0.04993803799152374
step: 110, loss: 0.07774066179990768
step: 120, loss: 0.23000861704349518
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.07503759860992432
step: 140, loss: 0.05210302397608757
step: 150, loss: 0.06270758807659149
step: 160, loss: 0.08329840749502182
step: 170, loss: 0.22118304669857025
step: 180, loss: 0.13705632090568542
step: 190, loss: 0.02379373461008072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7555555555555554, f1=0.7640449438202247,
best_f1=0.7640449438202247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.07992985844612122
step: 10, loss: 0.43324223160743713
step: 20, loss: 0.04929819330573082
step: 30, loss: 0.05113925039768219
step: 40, loss: 0.08200979232788086
step: 50, loss: 0.1977902501821518
step: 60, loss: 0.038000885397195816
step: 70, loss: 0.05825965851545334
step: 80, loss: 0.07461850345134735
step: 90, loss: 0.03749346733093262
step: 100, loss: 0.07456356287002563
step: 110, loss: 0.00846210028976202
step: 120, loss: 0.12033681571483612
step: 130, loss: 0.022306060418486595
step: 140, loss: 0.0912768766283989
step: 150, loss: 0.13073164224624634
step: 160, loss: 0.06551990658044815
step: 170, loss: 0.15912207961082458
step: 180, loss: 0.07141280919313431
step: 190, loss: 0.11989563703536987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7886597938144331, f1=0.7866323907455013,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017249418422579765
step: 10, loss: 0.15517771244049072
step: 20, loss: 0.0696665346622467
step: 30, loss: 0.03353990614414215
step: 40, loss: 0.0576777458190918
step: 50, loss: 0.02438347600400448
step: 60, loss: 0.06361372023820877
step: 70, loss: 0.07138462364673615
step: 80, loss: 0.14964933693408966
step: 90, loss: 0.07679935544729233
step: 100, loss: 0.1015840619802475
step: 110, loss: 0.004514442291110754
step: 120, loss: 0.010771763511002064
step: 130, loss: 0.1592855602502823
step: 140, loss: 0.033370569348335266
step: 150, loss: 0.023414311930537224

```



```

step: 160, loss: 0.004448088351637125
step: 170, loss: 0.14139840006828308
step: 180, loss: 0.01487011183053255
step: 190, loss: 0.10956712067127228
/usr/local/lib/python3.7/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=0.774928774928775, f1=0.7723342939481268,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03196394443511963
step: 10, loss: 0.02474421262741089
step: 20, loss: 0.016728904098272324
step: 30, loss: 0.02388031594455242
step: 40, loss: 0.05532189831137657
step: 50, loss: 0.0501253567636013
step: 60, loss: 0.07423042505979538
step: 70, loss: 0.04300284758210182
step: 80, loss: 0.003992130048573017
step: 90, loss: 0.007505107671022415
step: 100, loss: 0.0014418904902413487
step: 110, loss: 0.04299510270357132
step: 120, loss: 0.01913406513631344
step: 130, loss: 0.06328725814819336
step: 140, loss: 0.25764715671539307
step: 150, loss: 0.08898075670003891
step: 160, loss: 0.05742768198251724
step: 170, loss: 0.008280784823000431
step: 180, loss: 0.07629795372486115
step: 190, loss: 0.14670498669147491
/usr/local/lib/python3.7/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=0.7688022284122563, f1=0.779291553133515,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01897311769425869
step: 10, loss: 0.042571499943733215
step: 20, loss: 0.004247233271598816
step: 30, loss: 0.001025474863126874
step: 40, loss: 0.05519471317529678
step: 50, loss: 0.0027503662277013063
step: 60, loss: 0.021199170500040054
step: 70, loss: 0.006954134441912174

```

```

step: 80, loss: 0.14093990623950958
step: 90, loss: 0.009986682794988155
step: 100, loss: 0.0015236367471516132
step: 110, loss: 0.040379080921411514
step: 120, loss: 0.11835753917694092
step: 130, loss: 0.008231567218899727
step: 140, loss: 0.04322163015604019
step: 150, loss: 0.0005901755066588521
step: 160, loss: 0.09247776865959167
step: 170, loss: 0.018745770677924156
step: 180, loss: 0.008412241004407406
step: 190, loss: 0.044822826981544495
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.7945945945945946,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07919499278068542
step: 10, loss: 0.004700180608779192
step: 20, loss: 0.021742379292845726
step: 30, loss: 0.029057366773486137
step: 40, loss: 0.02020956762135029
step: 50, loss: 0.026214025914669037
step: 60, loss: 0.013125911355018616
step: 70, loss: 0.017268233001232147
step: 80, loss: 0.03236743435263634
step: 90, loss: 0.004422977101057768
step: 100, loss: 0.0010484647937119007
step: 110, loss: 0.09941588342189789
step: 120, loss: 0.0005682189366780221
step: 130, loss: 0.01793542504310608
step: 140, loss: 0.008780605159699917
step: 150, loss: 0.006366129033267498
step: 160, loss: 0.23340819776058197
step: 170, loss: 0.014387799426913261
step: 180, loss: 0.002502893563359976
step: 190, loss: 0.0009068108047358692
/usr/local/lib/python3.7/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=0.775623268698061, f1=0.7787114845938375,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.004579239059239626
step: 10, loss: 0.03219698369503021
step: 20, loss: 0.0010156575590372086
step: 30, loss: 0.012848976999521255
step: 40, loss: 0.0005746150854974985
step: 50, loss: 0.0004254999221302569
step: 60, loss: 0.0010913178557530046
step: 70, loss: 0.015240450389683247
step: 80, loss: 0.00043156472383998334
step: 90, loss: 0.00153925025369972
step: 100, loss: 0.03274841234087944
step: 110, loss: 0.0004552188911475241
step: 120, loss: 0.029855478554964066
step: 130, loss: 0.034181784838438034
step: 140, loss: 0.013530885800719261
step: 150, loss: 0.0018313780892640352
step: 160, loss: 0.0024259008932858706
step: 170, loss: 0.0010478494223207235
step: 180, loss: 0.007856632582843304
step: 190, loss: 0.0015830665361136198
/usr/local/lib/python3.7/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=0.7659574468085106, f1=0.75, best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011728830868378282
step: 10, loss: 0.01013440452516079
step: 20, loss: 0.0011758495820686221
step: 30, loss: 0.02890409156680107
step: 40, loss: 0.0008134891395457089
step: 50, loss: 0.0885097086429596
step: 60, loss: 0.0003066890931222588
step: 70, loss: 0.0005506501183845103
step: 80, loss: 0.0010896848980337381
step: 90, loss: 0.008180205710232258
step: 100, loss: 0.06992655247449875
step: 110, loss: 0.0006290447781793773
step: 120, loss: 0.007190283387899399
step: 130, loss: 0.023663299158215523
step: 140, loss: 0.011318703182041645
step: 150, loss: 0.003028024220839143
step: 160, loss: 0.0034624908585101366
step: 170, loss: 0.003847888670861721
step: 180, loss: 0.009531351737678051
step: 190, loss: 0.004902813117951155
/usr/local/lib/python3.7/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=0.7692307692307693, f1=0.7376237623762376,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010005056392401457
step: 10, loss: 0.000503475428558886
step: 20, loss: 0.030463000759482384
step: 30, loss: 0.0002817680360749364
step: 40, loss: 0.009275322780013084
step: 50, loss: 0.0003454362158663571
step: 60, loss: 0.001831082277931273
step: 70, loss: 0.0002780810173135251
step: 80, loss: 0.004700541030615568
step: 90, loss: 0.0002502446004655212
step: 100, loss: 0.00040056463330984116
step: 110, loss: 0.0003801570273935795
step: 120, loss: 0.0853923037648201
step: 130, loss: 0.0006305378046818078
step: 140, loss: 0.0017163835000246763
step: 150, loss: 0.014021406881511211
step: 160, loss: 0.0009746314026415348
step: 170, loss: 0.00015587407688144594
step: 180, loss: 0.0004098813224118203
step: 190, loss: 0.0004964368417859077
/usr/local/lib/python3.7/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=0.7713498622589531, f1=0.7891891891891892,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001758570724632591
step: 10, loss: 0.0017984795849770308
step: 20, loss: 0.0026852255687117577
step: 30, loss: 0.1257457137107849
step: 40, loss: 0.01699141599237919
step: 50, loss: 0.012422097846865654
step: 60, loss: 0.00019943006918765604
step: 70, loss: 0.000281556393019855
step: 80, loss: 0.0012978460872545838
step: 90, loss: 0.00025106663815677166
step: 100, loss: 0.0004915648605674505
step: 110, loss: 0.0010203705169260502
step: 120, loss: 0.0008439216762781143

```

```

step: 130, loss: 0.0003832708462141454
step: 140, loss: 0.00036544204340316355
step: 150, loss: 0.00038525561103597283
step: 160, loss: 0.001349756377749145
step: 170, loss: 0.0019001850159838796
step: 180, loss: 0.005805674009025097
step: 190, loss: 0.0005305922823026776
/usr/local/lib/python3.7/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=0.7774798927613942, f1=0.7598944591029023,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019309301569592208
step: 10, loss: 0.0016575738554820418
step: 20, loss: 0.00025862790062092245
step: 30, loss: 0.0014793070731684566
step: 40, loss: 0.0008243114571087062
step: 50, loss: 0.0008721117628738284
step: 60, loss: 0.0002045123401330784
step: 70, loss: 0.00011643006291706115
step: 80, loss: 0.00037790328497067094
step: 90, loss: 0.00020134566875640303
step: 100, loss: 0.0002644006453920156
step: 110, loss: 0.0014026073040440679
step: 120, loss: 0.0004104210820514709
step: 130, loss: 0.0005011525354348123
step: 140, loss: 0.0004343574109952897
step: 150, loss: 0.00022306967002805322
step: 160, loss: 0.00011043193080695346
step: 170, loss: 0.0004926965339109302
step: 180, loss: 0.013831213116645813
step: 190, loss: 0.0002484889992047101
/usr/local/lib/python3.7/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=0.7823691460055096, f1=0.7771739130434783,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011882680701091886
step: 10, loss: 0.0005756727186962962
step: 20, loss: 0.0005554115632548928
step: 30, loss: 0.006083985790610313
step: 40, loss: 0.005479574203491211

```

```

step: 50, loss: 0.000998387928120792
step: 60, loss: 0.0177077017724514
step: 70, loss: 0.005106587428599596
step: 80, loss: 0.0005417079082690179
step: 90, loss: 0.011362790130078793
step: 100, loss: 0.0008246580837294459
step: 110, loss: 0.0002577122941147536
step: 120, loss: 0.019587749615311623
step: 130, loss: 0.00022933611762709916
step: 140, loss: 0.000372821232303977
step: 150, loss: 0.00018788878514897078
step: 160, loss: 0.0031551963184028864
step: 170, loss: 0.0002847604628186673
step: 180, loss: 0.000687801802996546
step: 190, loss: 0.0026926430873572826
/usr/local/lib/python3.7/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=0.7811634349030471, f1=0.767123287671233,
best_f1=0.7866323907455013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010784219484776258
step: 10, loss: 0.0012112583499401808
step: 20, loss: 0.00046226728591136634
step: 30, loss: 0.0002226904616691172
step: 40, loss: 0.000407420244300738
step: 50, loss: 0.0002479505492374301
step: 60, loss: 0.00021285726688802242
step: 70, loss: 0.00038495211629197
step: 80, loss: 0.000191289305803366
step: 90, loss: 0.0009554795105941594
step: 100, loss: 0.0017358933109790087
step: 110, loss: 0.00030485031311400235
step: 120, loss: 0.0004397139127831906
step: 130, loss: 0.0017229776130989194
step: 140, loss: 0.0003275040362495929
step: 150, loss: 0.0013252852950245142
step: 160, loss: 0.0012257869821041822
step: 170, loss: 0.00019487563986331224
step: 180, loss: 0.00030726767727173865
step: 190, loss: 0.00025244790595024824
/usr/local/lib/python3.7/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=0.7816711590296496, f1=0.7634408602150539,
best_f1=0.7866323907455013

```

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

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

```
step: 0, loss: 0.0012836370151489973
step: 10, loss: 0.0003490015515126288
step: 20, loss: 0.0002616712008602917
step: 30, loss: 0.00023325403162743896
step: 40, loss: 0.0020399861969053745
step: 50, loss: 0.00021905545145273209
step: 60, loss: 0.00031067203963175416
step: 70, loss: 0.005290337838232517
step: 80, loss: 0.00047195740626193583
step: 90, loss: 0.00019680534023791552
step: 100, loss: 0.0002452176413498819
step: 110, loss: 0.0005350472638383508
step: 120, loss: 0.00012041180161759257
step: 130, loss: 0.0001912041479954496
step: 140, loss: 0.002870650729164481
step: 150, loss: 0.0002715869341045618
step: 160, loss: 0.00026656195404939353
step: 170, loss: 0.0010525741381570697
step: 180, loss: 0.003068910911679268
step: 190, loss: 0.00016802795289549977
```

```
/usr/local/lib/python3.7/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=0.7802197802197802, f1=0.7666666666666667,
best_f1=0.7866323907455013
```

```
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, 209.18it/s]
load_f1 = 0.7946666666666666
real_f1 = 0.7777777777777779
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 239.19it/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 bert-base-uncased \
    --fp16 \
    --save_model

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

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

```



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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

An input tensor was not cuda.

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

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

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

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

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

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

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

```
step: 0, loss: 0.6239549517631531  
step: 10, loss: 0.3644927740097046  
step: 20, loss: 0.30214232206344604  
step: 30, loss: 0.3997020721435547  
step: 40, loss: 0.26489588618278503  
step: 50, loss: 0.2506517767906189  
step: 60, loss: 0.22562791407108307
```

```

step: 70, loss: 0.3696611821651459
step: 80, loss: 0.3674132525920868
step: 90, loss: 0.27891528606414795
step: 100, loss: 0.1764945387840271
step: 110, loss: 0.2597387731075287
step: 120, loss: 0.14968127012252808
step: 130, loss: 0.09829188138246536
step: 140, loss: 0.21552544832229614
step: 150, loss: 0.23458747565746307
step: 160, loss: 0.11049719899892807
step: 170, loss: 0.15953540802001953
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.684931506849315, f1=0.6767241379310344,
best_f1=0.6767241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1451215147972107
step: 10, loss: 0.14293384552001953
step: 20, loss: 0.10058265924453735
step: 30, loss: 0.14649800956249237
step: 40, loss: 0.09313444048166275
step: 50, loss: 0.058134179562330246
step: 60, loss: 0.050605207681655884
step: 70, loss: 0.08027654886245728
step: 80, loss: 0.08088816702365875
step: 90, loss: 0.1444360613822937
step: 100, loss: 0.10278458893299103
step: 110, loss: 0.11600180715322495
step: 120, loss: 0.1466180980205536
step: 130, loss: 0.10510729253292084
step: 140, loss: 0.2612307369709015
step: 150, loss: 0.19447028636932373
step: 160, loss: 0.21321925520896912
step: 170, loss: 0.08046455681324005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7835990888382688, f1=0.7510729613733904,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04006575420498848
step: 10, loss: 0.08787377178668976
step: 20, loss: 0.08737371116876602
step: 30, loss: 0.07527664303779602
step: 40, loss: 0.062338147312402725
step: 50, loss: 0.1153559610247612
step: 60, loss: 0.09459174424409866
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.06669942289590836
step: 80, loss: 0.15419356524944305
step: 90, loss: 0.15589554607868195
step: 100, loss: 0.005847998894751072
step: 110, loss: 0.04934478923678398
step: 120, loss: 0.06882389634847641
step: 130, loss: 0.09542881697416306
step: 140, loss: 0.047125671058893204
step: 150, loss: 0.008319488726556301
step: 160, loss: 0.009993528015911579
step: 170, loss: 0.10938740521669388
/usr/local/lib/python3.7/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=0.7655502392344498, f1=0.7654867256637168,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006681045051664114
step: 10, loss: 0.050848282873630524
step: 20, loss: 0.004457369446754456
step: 30, loss: 0.05705980584025383
step: 40, loss: 0.003186634974554181
step: 50, loss: 0.15205752849578857
step: 60, loss: 0.15774665772914886
step: 70, loss: 0.003230802249163389
step: 80, loss: 0.15056756138801575
step: 90, loss: 0.08277793228626251
step: 100, loss: 0.2086234837770462
step: 110, loss: 0.10318002104759216
step: 120, loss: 0.014814703725278378
step: 130, loss: 0.1277744621038437
step: 140, loss: 0.06744074821472168
step: 150, loss: 0.11726093292236328

```

```

step: 160, loss: 0.17072957754135132
step: 170, loss: 0.006282032933086157
/usr/local/lib/python3.7/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=0.7453083109919573, f1=0.7707317073170732,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012058100663125515
step: 10, loss: 0.012293625622987747
step: 20, loss: 0.027362829074263573
step: 30, loss: 0.014387885108590126
step: 40, loss: 0.027826180681586266
step: 50, loss: 0.03178347647190094
step: 60, loss: 0.019326284527778625
step: 70, loss: 0.2096254974603653
step: 80, loss: 0.018761150538921356
step: 90, loss: 0.04479164630174637
step: 100, loss: 0.015542982146143913
step: 110, loss: 0.05260763689875603
step: 120, loss: 0.018492573872208595
step: 130, loss: 0.011384237557649612
step: 140, loss: 0.0022535850293934345
step: 150, loss: 0.027067633345723152
step: 160, loss: 0.015577939338982105
step: 170, loss: 0.014140523038804531
/usr/local/lib/python3.7/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=0.7769784172661871, f1=0.7534246575342466,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005783699452877045
step: 10, loss: 0.01389599684625864
step: 20, loss: 0.0034332536160945892
step: 30, loss: 0.021856367588043213
step: 40, loss: 0.008592902682721615
step: 50, loss: 0.17913630604743958
step: 60, loss: 0.03390977159142494
step: 70, loss: 0.0512935109436512
step: 80, loss: 0.019408589228987694
step: 90, loss: 0.11570417135953903
step: 100, loss: 0.0018826484447345138
step: 110, loss: 0.0005786063265986741

```

```

step: 120, loss: 0.05025705695152283
step: 130, loss: 0.007128555793315172
step: 140, loss: 0.02321595326066017
step: 150, loss: 0.06823965162038803
step: 160, loss: 0.05934230238199234
step: 170, loss: 0.0019652184564620256
/usr/local/lib/python3.7/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=0.7627906976744185, f1=0.7415730337078651,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025395057164132595
step: 10, loss: 0.017733603715896606
step: 20, loss: 0.0018768924055621028
step: 30, loss: 0.06047782674431801
step: 40, loss: 0.0009911232627928257
step: 50, loss: 0.03361676633358002
step: 60, loss: 0.03079228475689888
step: 70, loss: 0.0007339478470385075
step: 80, loss: 0.011775581166148186
step: 90, loss: 0.00022394749976228923
step: 100, loss: 0.030174318701028824
step: 110, loss: 0.07286137342453003
step: 120, loss: 0.002934057265520096
step: 130, loss: 0.157416433095932
step: 140, loss: 0.002398180775344372
step: 150, loss: 0.004468163009732962
step: 160, loss: 0.00031072806450538337
step: 170, loss: 0.09087162464857101
/usr/local/lib/python3.7/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=0.7397959183673468, f1=0.7584541062801932,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001031082239933312
step: 10, loss: 0.007406397257000208
step: 20, loss: 0.0005862468387931585
step: 30, loss: 0.023381059989333153
step: 40, loss: 0.00020762377243954688
step: 50, loss: 0.002631204901263118
step: 60, loss: 0.0006498451693914831
step: 70, loss: 0.00048189147491939366

```

```

step: 80, loss: 0.012126632034778595
step: 90, loss: 0.00048116949619725347
step: 100, loss: 0.022617720067501068
step: 110, loss: 0.0777871385216713
step: 120, loss: 0.028559472411870956
step: 130, loss: 0.06412993371486664
step: 140, loss: 0.018508540466427803
step: 150, loss: 0.007007207255810499
step: 160, loss: 0.001055861241184175
step: 170, loss: 0.012382525019347668
/usr/local/lib/python3.7/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=0.7524752475247525, f1=0.7777777777777778,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008596485713496804
step: 10, loss: 0.002945365384221077
step: 20, loss: 0.04880654811859131
step: 30, loss: 0.02890835702419281
step: 40, loss: 0.006979696452617645
step: 50, loss: 0.000398524891352281
step: 60, loss: 0.0047696176916360855
step: 70, loss: 0.004246257245540619
step: 80, loss: 0.043466053903102875
step: 90, loss: 0.030970225110650063
step: 100, loss: 0.00739689264446497
step: 110, loss: 0.004372366238385439
step: 120, loss: 0.24745069444179535
step: 130, loss: 0.04626578092575073
step: 140, loss: 0.0011587456101551652
step: 150, loss: 0.01583675481379032
step: 160, loss: 0.049188282340765
step: 170, loss: 0.08560748398303986
/usr/local/lib/python3.7/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=0.7656250000000001, f1=0.7949367088607596,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.042278390377759933
step: 10, loss: 0.0015937520656734705
step: 20, loss: 0.07657880336046219
step: 30, loss: 0.04153688997030258

```

```

step: 40, loss: 0.00016561477968934923
step: 50, loss: 0.08668022602796555
step: 60, loss: 0.00014226773055270314
step: 70, loss: 0.0008866913849487901
step: 80, loss: 0.0022066538222134113
step: 90, loss: 0.05454636365175247
step: 100, loss: 0.00018450099742040038
step: 110, loss: 0.0011581386206671596
step: 120, loss: 0.0002827552380040288
step: 130, loss: 0.016167039051651955
step: 140, loss: 0.010589847341179848
step: 150, loss: 0.002683401107788086
step: 160, loss: 0.013200144283473492
step: 170, loss: 0.00020466066780500114
/usr/local/lib/python3.7/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=0.7291666666666666, f1=0.7696078431372548,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045927561819553375
step: 10, loss: 0.03071884624660015
step: 20, loss: 0.0027519981376826763
step: 30, loss: 0.0001070531434379518
step: 40, loss: 0.00041670806240290403
step: 50, loss: 0.0021622313652187586
step: 60, loss: 0.0182994082570076
step: 70, loss: 0.000121540084364824
step: 80, loss: 0.005538585130125284
step: 90, loss: 0.0001318944268859923
step: 100, loss: 8.304806397063658e-05
step: 110, loss: 0.024504059925675392
step: 120, loss: 0.00010377904254710302
step: 130, loss: 0.00011308120883768424
step: 140, loss: 0.00019776758563239127
step: 150, loss: 0.015002504922449589
step: 160, loss: 0.01481225248426199
step: 170, loss: 0.046103257685899734
/usr/local/lib/python3.7/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=0.7539503386004516, f1=0.7586206896551724,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0013433507410809398
step: 10, loss: 0.0006642616353929043
step: 20, loss: 0.0016270079649984837
step: 30, loss: 0.00026027802960015833
step: 40, loss: 0.0004947073175571859
step: 50, loss: 0.00010084340465255082
step: 60, loss: 0.005247986409813166
step: 70, loss: 0.09242638200521469
step: 80, loss: 0.0002325812674826011
step: 90, loss: 0.001280836295336485
step: 100, loss: 0.03796524554491043
step: 110, loss: 0.00022199000522959977
step: 120, loss: 0.004735978785902262
step: 130, loss: 0.031089385971426964
step: 140, loss: 0.002450898289680481
step: 150, loss: 0.015674948692321777
step: 160, loss: 8.698267629370093e-05
step: 170, loss: 0.00010306775948265567
/usr/local/lib/python3.7/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=0.75, f1=0.7794117647058824, best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03027370013296604
step: 10, loss: 0.00010896843014052138
step: 20, loss: 0.002568761119619012
step: 30, loss: 0.00011018141231033951
step: 40, loss: 0.00016988808056339622
step: 50, loss: 0.00021583469060715288
step: 60, loss: 0.00970734003931284
step: 70, loss: 9.631055581849068e-05
step: 80, loss: 0.00026792005519382656
step: 90, loss: 7.205089059425518e-05
step: 100, loss: 0.00018576164438854903
step: 110, loss: 7.639001705683768e-05
step: 120, loss: 0.040617480874061584
step: 130, loss: 0.00013341347221285105
step: 140, loss: 0.0019481431227177382
step: 150, loss: 0.011277720332145691
step: 160, loss: 0.01241240929812193
step: 170, loss: 0.00037253391928970814
/usr/local/lib/python3.7/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=0.738544474393531, f1=0.7777777777777778,
best_f1=0.7510729613733904

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026054243789985776
step: 10, loss: 0.0002165762853110209
step: 20, loss: 0.0005199161823838949
step: 30, loss: 0.000492817722260952
step: 40, loss: 9.70285982475616e-05
step: 50, loss: 0.00014710989489685744
step: 60, loss: 6.157141615403816e-05
step: 70, loss: 0.00010017097520176321
step: 80, loss: 0.019103430211544037
step: 90, loss: 7.143385300878435e-05
step: 100, loss: 7.168607407948002e-05
step: 110, loss: 0.00013577926438301802
step: 120, loss: 0.02873814105987549
step: 130, loss: 0.0015185244847089052
step: 140, loss: 0.00021152853150852025
step: 150, loss: 5.696331209037453e-05
step: 160, loss: 4.140556848142296e-05
step: 170, loss: 0.00023355762823484838
/usr/local/lib/python3.7/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=0.7468671679197996, f1=0.7743467933491687,
best_f1=0.7510729613733904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019623126718215644
step: 10, loss: 0.004982075188308954
step: 20, loss: 0.00481564411893487
step: 30, loss: 0.00011970914783887565
step: 40, loss: 0.0016024982323870063
step: 50, loss: 0.0003056882997043431
step: 60, loss: 0.009810992516577244
step: 70, loss: 6.756679795216769e-05
step: 80, loss: 0.029795264825224876
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.022275423631072044
step: 100, loss: 0.0017467443831264973
step: 110, loss: 0.00011762812937377021
step: 120, loss: 0.00028073618886992335
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.00639710808172822
step: 140, loss: 0.00021377996017690748
step: 150, loss: 0.00030744788818992674
step: 160, loss: 7.474834274034947e-05

```

```

step: 170, loss: 8.776242611929774e-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=0.7424242424242425, f1=0.7769784172661871,
best_f1=0.7510729613733904
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:07, 258.40it/s]
load_f1 = 0.782608695652174
real_f1 = 0.7835990888382688
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:19, 223.37it/s]

```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

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

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

Downloading: 100% 433/433 [00:00<00:00, 393kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 254kB/s]

Downloading: 100% 440M/440M [00:09<00:00, 47.9MB/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: 0.6112598180770874
step: 10, loss: 0.6152108907699585
step: 20, loss: 0.4864363372325897
step: 30, loss: 0.38060060143470764
step: 40, loss: 0.094855397939682
step: 50, loss: 0.06265205889940262
step: 60, loss: 0.11542122811079025
step: 70, loss: 0.32022228837013245
step: 80, loss: 0.1829175353050232
step: 90, loss: 0.11884397268295288
step: 100, loss: 0.005959603004157543
step: 110, loss: 0.2145020067691803
step: 120, loss: 0.02188260108232498
step: 130, loss: 0.022124789655208588
step: 140, loss: 0.0024032953660935163
step: 150, loss: 0.04651360958814621
step: 160, loss: 0.004309630021452904
step: 170, loss: 0.17877928912639618
step: 180, loss: 0.12943296134471893
step: 190, loss: 0.097454734146595
step: 200, loss: 0.07755225151777267
step: 210, loss: 0.0068781436420977116
step: 220, loss: 0.02093317173421383
step: 230, loss: 0.012132402509450912
```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9696287964004499, f1=0.9661399548532732,
best_f1=0.9661399548532732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017761055380105972
step: 10, loss: 0.03677797317504883
step: 20, loss: 0.16007240116596222
step: 30, loss: 0.1905786246061325
step: 40, loss: 0.16851624846458435
step: 50, loss: 0.005362022668123245
step: 60, loss: 0.005014074966311455
step: 70, loss: 0.12689365446567535
step: 80, loss: 0.03522759675979614
step: 90, loss: 0.04147486388683319
step: 100, loss: 0.07543858140707016
step: 110, loss: 0.10163252800703049
step: 120, loss: 0.2136780172586441
step: 130, loss: 0.04356838017702103
step: 140, loss: 0.00261681922711432
step: 150, loss: 0.007440528832376003
step: 160, loss: 0.01776772364974022
step: 170, loss: 0.0021364353597164154
step: 180, loss: 0.011055308394134045
step: 190, loss: 0.050127461552619934
step: 200, loss: 0.002023361623287201
step: 210, loss: 0.0025923997163772583
step: 220, loss: 0.18404290080070496
step: 230, loss: 0.10068903863430023
/usr/local/lib/python3.7/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=0.9669327251995439, f1=0.956120092378753,
best_f1=0.9661399548532732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02152126468718052
step: 10, loss: 0.016559679061174393
step: 20, loss: 0.009016918949782848
step: 30, loss: 0.03335751220583916

```

```

step: 40, loss: 0.05995248630642891
step: 50, loss: 0.004714990966022015
step: 60, loss: 0.0059843952767550945
step: 70, loss: 0.002693263115361333
step: 80, loss: 0.002910967916250229
step: 90, loss: 0.17314985394477844
step: 100, loss: 0.007223344873636961
step: 110, loss: 0.0008886882569640875
step: 120, loss: 0.0040726978331804276
step: 130, loss: 0.0009380722767673433
step: 140, loss: 0.010433640331029892
step: 150, loss: 0.011056297458708286
step: 160, loss: 0.06320201605558395
step: 170, loss: 0.004290483891963959
step: 180, loss: 0.012245126999914646
step: 190, loss: 0.006676229182630777
step: 200, loss: 0.004117279313504696
step: 210, loss: 0.0028935777954757214
step: 220, loss: 0.006898805033415556
step: 230, loss: 0.003089780453592539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9700996677740864, 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: 0.027439473196864128
step: 10, loss: 0.003451394382864237
step: 20, loss: 0.0009797634556889534
step: 30, loss: 0.0013335032854229212
step: 40, loss: 0.00886398646980524
step: 50, loss: 0.002304572146385908
step: 60, loss: 0.005714309401810169
step: 70, loss: 0.0010186698054894805
step: 80, loss: 0.002255622297525406
step: 90, loss: 0.03064051643013954
step: 100, loss: 0.003908065613359213
step: 110, loss: 0.0009464774047955871
step: 120, loss: 0.0048616728745400906
step: 130, loss: 0.0022548683919012547
step: 140, loss: 0.0027905094902962446
step: 150, loss: 0.18327955901622772

```

```

step: 160, loss: 0.07993868738412857
step: 170, loss: 0.04136517643928528
step: 180, loss: 0.0006748127634637058
step: 190, loss: 0.006856371648609638
step: 200, loss: 0.001362305716611445
step: 210, loss: 0.050576046109199524
step: 220, loss: 0.0004819448513444513
step: 230, loss: 0.03286813944578171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9754464285714286, 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: 0.004111513029783964
step: 10, loss: 0.002645275089889765
step: 20, loss: 0.008470257744193077
step: 30, loss: 0.0005782552761957049
step: 40, loss: 0.0008293294231407344
step: 50, loss: 0.0012007192708551884
step: 60, loss: 0.06629544496536255
step: 70, loss: 0.0008569030906073749
step: 80, loss: 0.001277924864552915
step: 90, loss: 0.014365624636411667
step: 100, loss: 0.001667998032644391
step: 110, loss: 0.0008569390629418194
step: 120, loss: 0.0002553225785959512
step: 130, loss: 0.0021814105566591024
step: 140, loss: 0.0006903365720063448
step: 150, loss: 0.0007551733870059252
step: 160, loss: 0.009164618328213692
step: 170, loss: 0.00686703622341156
step: 180, loss: 0.0011860784143209457
step: 190, loss: 0.08534671366214752
step: 200, loss: 0.01090079452842474
step: 210, loss: 0.004904784262180328
step: 220, loss: 0.002343252766877413
step: 230, loss: 0.0008954696240834892
/usr/local/lib/python3.7/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=0.9730941704035874, f1=0.9731543624161074,

```

```

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: 0.004779621493071318
step: 10, loss: 0.0015246541006490588
step: 20, loss: 0.0010965915862470865
step: 30, loss: 0.0003535410505719483
step: 40, loss: 0.0036727087572216988
step: 50, loss: 0.0018093979451805353
step: 60, loss: 0.0006231339648365974
step: 70, loss: 0.0011157778790220618
step: 80, loss: 0.0015597464516758919
step: 90, loss: 0.0010378953302279115
step: 100, loss: 0.023331711068749428
step: 110, loss: 0.0025808101054280996
step: 120, loss: 0.0008479421376250684
step: 130, loss: 0.0006550696562044322
step: 140, loss: 0.0002837730571627617
step: 150, loss: 0.0018815535586327314
step: 160, loss: 0.009572270326316357
step: 170, loss: 0.00025002870825119317
step: 180, loss: 0.0021588595118373632
step: 190, loss: 0.00807722844183445
step: 200, loss: 0.0007635706569999456
step: 210, loss: 0.0002481662668287754
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.0002903466229327023
step: 230, loss: 0.029594015330076218
/usr/local/lib/python3.7/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=0.9670329670329669, f1=0.9700996677740864,
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: 0.0023523776326328516
step: 10, loss: 0.00039266340900212526
step: 20, loss: 0.0014777987962588668
step: 30, loss: 0.0011990144848823547
step: 40, loss: 0.0049383025616407394
step: 50, loss: 0.0004961728118360043
step: 60, loss: 0.0009196443133987486
step: 70, loss: 0.029861856251955032
step: 80, loss: 0.002409969223663211
step: 90, loss: 0.0008273026323877275
step: 100, loss: 0.00022901604825165123

```



```

step: 110, loss: 0.00021636085875798017
step: 120, loss: 0.00015396713570225984
step: 130, loss: 0.0004282438021618873
step: 140, loss: 0.00036329225986264646
step: 150, loss: 0.0031336224637925625
step: 160, loss: 0.027230767533183098
step: 170, loss: 0.0010874946601688862
step: 180, loss: 0.0025195125490427017
step: 190, loss: 0.0011221232125535607
step: 200, loss: 0.040697358548641205
step: 210, loss: 0.00018256386101711541
step: 220, loss: 0.002101701218634844
step: 230, loss: 0.0012492014793679118
/usr/local/lib/python3.7/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=0.9741282339707535, f1=0.9684684684684683,
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: 0.003957483917474747
step: 10, loss: 0.0015444968594238162
step: 20, loss: 0.022057076916098595
step: 30, loss: 0.00037131423596292734
step: 40, loss: 0.002222563372924924
step: 50, loss: 0.0003025274781975895
step: 60, loss: 0.0008862779941409826
step: 70, loss: 0.00034274120116606355
step: 80, loss: 0.0026965527795255184
step: 90, loss: 0.0003145029186271131
step: 100, loss: 0.0005428788135759532
step: 110, loss: 0.0019082234939560294
step: 120, loss: 0.06502757966518402
step: 130, loss: 0.021146021783351898
step: 140, loss: 0.001999922562390566
step: 150, loss: 0.0009509600349701941
step: 160, loss: 0.0004479696508497
step: 170, loss: 0.002791621722280979
step: 180, loss: 0.00277327629737556
step: 190, loss: 0.0022725414019078016
step: 200, loss: 0.0015563148772343993
step: 210, loss: 0.005544724408537149
step: 220, loss: 0.0002717715688049793
step: 230, loss: 0.00021103133622091264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9764837625979844, f1=0.9743589743589743,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.616029324708506e-05
step: 10, loss: 0.0010417011799290776
step: 20, loss: 0.014735774137079716
step: 30, loss: 0.00018787947192322463
step: 40, loss: 0.12748901546001434
step: 50, loss: 7.535951590398327e-05
step: 60, loss: 0.00018251659639645368
step: 70, loss: 0.027294320985674858
step: 80, loss: 0.0013373931869864464
step: 90, loss: 0.0008228015503846109
step: 100, loss: 0.0009864743333309889
step: 110, loss: 8.522640564478934e-05
step: 120, loss: 0.0009766501607373357
step: 130, loss: 0.00014182455197442323
step: 140, loss: 9.015727846417576e-05
step: 150, loss: 0.0074680703692138195
step: 160, loss: 0.00010153687617275864
step: 170, loss: 0.00014916641521267593
step: 180, loss: 0.00019356027769390494
step: 190, loss: 0.00020429090363904834
step: 200, loss: 0.00032112569897435606
step: 210, loss: 0.0004586279974319041
step: 220, loss: 9.761336696101353e-05
step: 230, loss: 0.00013294158270582557
/usr/local/lib/python3.7/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=0.9754464285714286, f1=0.9732142857142857,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.893933747662231e-05
step: 10, loss: 9.267417044611648e-05
step: 20, loss: 6.729217420797795e-05
step: 30, loss: 8.892144251149148e-05
step: 40, loss: 8.518352115061134e-05
step: 50, loss: 0.0005029244348406792
step: 60, loss: 0.00010627121810102835

```

```

step: 70, loss: 8.865290874382481e-05
step: 80, loss: 9.805605077417567e-05
step: 90, loss: 0.00018592984997667372
step: 100, loss: 7.118645589798689e-05
step: 110, loss: 0.000230917488806881
step: 120, loss: 5.8493158576311544e-05
step: 130, loss: 0.0003354742075316608
step: 140, loss: 0.00149279716424644
step: 150, loss: 0.00011180900037288666
step: 160, loss: 4.239055488142185e-05
step: 170, loss: 6.134222348919138e-05
step: 180, loss: 0.0001260146964341402
step: 190, loss: 0.0006609742995351553
step: 200, loss: 5.891596811125055e-05
step: 210, loss: 0.00014090209151618183
step: 220, loss: 8.380268991459161e-05
step: 230, loss: 0.00040287189767695963
/usr/local/lib/python3.7/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=0.9743016759776536, f1=0.9754464285714286,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.233834574231878e-05
step: 10, loss: 7.647690654266626e-05
step: 20, loss: 0.002339928410947323
step: 30, loss: 0.013605077750980854
step: 40, loss: 0.00015194635489024222
step: 50, loss: 0.0002666124200914055
step: 60, loss: 0.0004372056864667684
step: 70, loss: 0.0001824582868721336
step: 80, loss: 6.892256351420656e-05
step: 90, loss: 0.001097531639970839
step: 100, loss: 0.04573452100157738
step: 110, loss: 0.00036603320040740073
step: 120, loss: 0.010354443453252316
step: 130, loss: 0.0003635360044427216
step: 140, loss: 6.973755807848647e-05
step: 150, loss: 0.003822747152298689
step: 160, loss: 8.395194163313136e-05
step: 170, loss: 0.021087417379021645
step: 180, loss: 0.0001654986699577421
step: 190, loss: 4.9218800995731726e-05
step: 200, loss: 7.279128476511687e-05
step: 210, loss: 0.00010209358879365027
step: 220, loss: 7.033345173113048e-05

```

```

step: 230, loss: 4.916274338029325e-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=0.9741282339707535, f1=0.9752252252252253,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.590608194936067e-05
step: 10, loss: 4.818597881239839e-05
step: 20, loss: 7.05946731613949e-05
step: 30, loss: 0.00016988973948173225
step: 40, loss: 0.0006815597880631685
step: 50, loss: 8.508073369739577e-05
step: 60, loss: 0.00021239712077658623
step: 70, loss: 7.40468967705965e-05
step: 80, loss: 9.7717231255956e-05
step: 90, loss: 5.429416705737822e-05
step: 100, loss: 4.0267957956530154e-05
step: 110, loss: 7.74035943322815e-05
step: 120, loss: 5.31918449269142e-05
step: 130, loss: 5.8224188251188025e-05
step: 140, loss: 5.8847035688813776e-05
step: 150, loss: 0.00010467137326486409
step: 160, loss: 8.241917385021225e-05
step: 170, loss: 4.974089824827388e-05
step: 180, loss: 0.0008077114471234381
step: 190, loss: 8.60097716213204e-05
step: 200, loss: 2.9737677323282696e-05
step: 210, loss: 0.00017541254055686295
step: 220, loss: 7.235787052195519e-05
step: 230, loss: 0.00027195014990866184
/usr/local/lib/python3.7/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=0.971815107102593, f1=0.9704545454545453,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.554352876264602e-05
step: 10, loss: 6.795408990001306e-05
step: 20, loss: 4.206008088658564e-05
step: 30, loss: 9.791870979825035e-05
step: 40, loss: 0.0009538052254356444
step: 50, loss: 0.00018215292948298156
step: 60, loss: 5.14877428940963e-05

```

```

step: 70, loss: 3.732525874511339e-05
step: 80, loss: 4.302862362237647e-05
step: 90, loss: 0.00010108546848641708
step: 100, loss: 3.3835487556643784e-05
step: 110, loss: 0.003525231033563614
step: 120, loss: 0.04291420057415962
step: 130, loss: 6.118675082689151e-05
step: 140, loss: 7.842678314773366e-05
step: 150, loss: 4.925293251289986e-05
step: 160, loss: 2.754759589151945e-05
step: 170, loss: 7.039880438242108e-05
step: 180, loss: 4.343091859482229e-05
step: 190, loss: 6.772933556931093e-05
step: 200, loss: 5.851779496879317e-05
step: 210, loss: 3.03709148283815e-05
step: 220, loss: 4.535993139143102e-05
step: 230, loss: 3.517985169310123e-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=0.9719416386083053, f1=0.9695603156708005,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.597681203857064e-05
step: 10, loss: 9.149573452305049e-05
step: 20, loss: 3.400663990760222e-05
step: 30, loss: 6.219232454895973e-05
step: 40, loss: 6.895814294693992e-05
step: 50, loss: 8.729423279874027e-05
step: 60, loss: 0.00012115918070776388
step: 70, loss: 5.0492759328335524e-05
step: 80, loss: 8.300030458485708e-05
step: 90, loss: 6.788083555875346e-05
step: 100, loss: 4.263335722498596e-05
step: 110, loss: 3.810392809100449e-05
step: 120, loss: 2.4001425117603503e-05
step: 130, loss: 4.220930713927373e-05
step: 140, loss: 5.0954815378645435e-05
step: 150, loss: 8.527628233423457e-05
step: 160, loss: 0.00024335639318451285
step: 170, loss: 2.036949445027858e-05
step: 180, loss: 6.360584666253999e-05
step: 190, loss: 5.68809628020972e-05
step: 200, loss: 3.878834831994027e-05
step: 210, loss: 4.2645180656109005e-05
step: 220, loss: 5.986088581266813e-05

```

```

step: 230, loss: 0.010231570340692997
/usr/local/lib/python3.7/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=0.9730337078651685, f1=0.9706546275395034,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.501586616039276e-05
step: 10, loss: 2.1967329303151928e-05
step: 20, loss: 4.4965490815229714e-05
step: 30, loss: 4.943025123793632e-05
step: 40, loss: 8.002802496775985e-05
step: 50, loss: 0.00018259287753608078
step: 60, loss: 0.00012974426499567926
step: 70, loss: 4.16946058976464e-05
step: 80, loss: 4.1090537706622854e-05
step: 90, loss: 9.970812243409455e-05
step: 100, loss: 6.053608376532793e-05
step: 110, loss: 3.801816637860611e-05
step: 120, loss: 8.413288742303848e-05
step: 130, loss: 8.134677773341537e-05
step: 140, loss: 4.052490112371743e-05
step: 150, loss: 0.00012177725875517353
step: 160, loss: 3.0121587769826874e-05
step: 170, loss: 3.322736301925033e-05
step: 180, loss: 0.00022261975391302258
step: 190, loss: 9.54874703893438e-05
step: 200, loss: 9.834737284108996e-05
step: 210, loss: 9.982440678868443e-05
step: 220, loss: 4.373166666482575e-05
step: 230, loss: 3.443870082264766e-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=0.9718785151856018, f1=0.9695603156708005,
best_f1=0.9743589743589743
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 [00:14, 168.61it/s]
load_f1 = 0.9752808988764046
real_f1 = 0.9752808988764046
0it [00:00, ?it/s]/usr/local/lib/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, 191.13it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

```

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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


UserWarning)

step: 0, loss: 0.6308393478393555
step: 10, loss: 0.5505378842353821
step: 20, loss: 0.5521015524864197
step: 30, loss: 0.1809433549642563
step: 40, loss: 0.17599156498908997
step: 50, loss: 0.2394127994775772
step: 60, loss: 0.18213963508605957
step: 70, loss: 0.14485208690166473
step: 80, loss: 0.023127535358071327
step: 90, loss: 0.44621583819389343
step: 100, loss: 0.05394165217876434
step: 110, loss: 0.14710931479930878
step: 120, loss: 0.1885993629693985
step: 130, loss: 0.05668529123067856
step: 140, loss: 0.11335673928260803
step: 150, loss: 0.04183400049805641
step: 160, loss: 0.03932279720902443
step: 170, loss: 0.11107391119003296
step: 180, loss: 0.04098522290587425
step: 190, loss: 0.011395781300961971
step: 200, loss: 0.2622057795524597
step: 210, loss: 0.0969039648771286
step: 220, loss: 0.22201977670192719
step: 230, loss: 0.17912285029888153
step: 240, loss: 0.09679540246725082
step: 250, loss: 0.02308247983455658
step: 260, loss: 0.15860846638679504
step: 270, loss: 0.05051843822002411
step: 280, loss: 0.06804773956537247
step: 290, loss: 0.06991148740053177
step: 300, loss: 0.034095924347639084
step: 310, loss: 0.23222658038139343
step: 320, loss: 0.10027280449867249
step: 330, loss: 0.029407698661088943
step: 340, loss: 0.0958208441734314
step: 350, loss: 0.13833585381507874
step: 360, loss: 0.06409045308828354
step: 370, loss: 0.09021957963705063
step: 380, loss: 0.022184494882822037
step: 390, loss: 0.17274907231330872
step: 400, loss: 0.2700599431991577
step: 410, loss: 0.051091961562633514
step: 420, loss: 0.038017839193344116
step: 430, loss: 0.1589588224887848
step: 440, loss: 0.022621162235736847
step: 450, loss: 0.013283144682645798
step: 460, loss: 0.01073786523193121

```

step: 470, loss: 0.16342288255691528
step: 480, loss: 0.09255707263946533
step: 490, loss: 0.05606403201818466
step: 500, loss: 0.12096572667360306
step: 510, loss: 0.09334856271743774
step: 520, loss: 0.04892009124159813
step: 530, loss: 0.008651681244373322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9247211895910781, f1=0.92201199815413,
best_f1=0.92201199815413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2184745967388153
step: 10, loss: 0.03361574560403824
step: 20, loss: 0.0132715729996562
step: 30, loss: 0.02078947052359581
step: 40, loss: 0.04732947424054146
step: 50, loss: 0.27160024642944336
step: 60, loss: 0.005717413034290075
step: 70, loss: 0.05986982211470604
step: 80, loss: 0.09475748986005783
step: 90, loss: 0.019000547006726265
step: 100, loss: 0.0746271088719368
step: 110, loss: 0.009987887926399708
step: 120, loss: 0.07617209106683731
step: 130, loss: 0.18185296654701233
step: 140, loss: 0.012068985030055046
step: 150, loss: 0.13188256323337555
step: 160, loss: 0.06945765763521194
step: 170, loss: 0.016961287707090378
step: 180, loss: 0.011363590136170387
step: 190, loss: 0.08062175661325455
step: 200, loss: 0.01122385822236538
step: 210, loss: 0.09369737654924393
step: 220, loss: 0.08627444505691528
step: 230, loss: 0.007848703302443027
step: 240, loss: 0.02608400210738182
step: 250, loss: 0.01381416991353035
step: 260, loss: 0.0033317916095256805
step: 270, loss: 0.27997711300849915
step: 280, loss: 0.10493452847003937

```

```

step: 290, loss: 0.012776427902281284
step: 300, loss: 0.1075260117650032
step: 310, loss: 0.03270461782813072
step: 320, loss: 0.14427253603935242
step: 330, loss: 0.06882183253765106
step: 340, loss: 0.12521225214004517
step: 350, loss: 0.003968583885580301
step: 360, loss: 0.15966898202896118
step: 370, loss: 0.14020118117332458
step: 380, loss: 0.06103042513132095
step: 390, loss: 0.05577564612030983
step: 400, loss: 0.11468195915222168
step: 410, loss: 0.027708174660801888
step: 420, loss: 0.11023019999265671
step: 430, loss: 0.017893552780151367
step: 440, loss: 0.1965186446905136
step: 450, loss: 0.012544283643364906
step: 460, loss: 0.13696977496147156
step: 470, loss: 0.05711296573281288
step: 480, loss: 0.24070660769939423
step: 490, loss: 0.05794903635978699
step: 500, loss: 0.35952097177505493
step: 510, loss: 0.04695383086800575
step: 520, loss: 0.058236394077539444
step: 530, loss: 0.02270045317709446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9255121042830541, f1=0.9255813953488372,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031204763799905777
step: 10, loss: 0.017232827842235565
step: 20, loss: 0.10460646450519562
step: 30, loss: 0.029482485726475716
step: 40, loss: 0.01494636945426464
step: 50, loss: 0.07312681525945663
step: 60, loss: 0.010425076819956303
step: 70, loss: 0.03702405467629433
step: 80, loss: 0.015442348085343838
step: 90, loss: 0.0034840202424675226
step: 100, loss: 0.1551879644393921

```

```
step: 110, loss: 0.038257189095020294
step: 120, loss: 0.01130492053925991
step: 130, loss: 0.0016040618065744638
step: 140, loss: 0.043392717838287354
step: 150, loss: 0.05259861797094345
step: 160, loss: 0.007635003421455622
step: 170, loss: 0.05638331174850464
step: 180, loss: 0.02072395756840706
step: 190, loss: 0.01383789349347353
step: 200, loss: 0.021570052951574326
step: 210, loss: 0.0768694058060646
step: 220, loss: 0.06791865825653076
step: 230, loss: 0.08925943076610565
step: 240, loss: 0.007134330924600363
step: 250, loss: 0.07419874519109726
step: 260, loss: 0.03575114160776138
step: 270, loss: 0.012243632227182388
step: 280, loss: 0.02952844649553299
step: 290, loss: 0.002789998659864068
step: 300, loss: 0.009110583923757076
step: 310, loss: 0.005960981361567974
step: 320, loss: 0.01758362352848053
step: 330, loss: 0.0014191899681463838
step: 340, loss: 0.0011563734151422977
step: 350, loss: 0.007553056348115206
step: 360, loss: 0.07608353346586227
step: 370, loss: 0.011820752173662186
step: 380, loss: 0.056717973202466965
step: 390, loss: 0.06898420304059982
step: 400, loss: 0.06119031459093094
step: 410, loss: 0.01774400845170021
step: 420, loss: 0.2998906075954437
step: 430, loss: 0.022221412509679794
step: 440, loss: 0.008416597731411457
step: 450, loss: 0.07857640087604523
step: 460, loss: 0.02338864468038082
step: 470, loss: 0.01630321517586708
step: 480, loss: 0.0025798112619668245
step: 490, loss: 0.002977826865389943
step: 500, loss: 0.002186661120504141
step: 510, loss: 0.0021563328336924314
step: 520, loss: 0.10120730847120285
step: 530, loss: 0.01043766364455223
/usr/local/lib/python3.7/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=0.9177767570050528, f1=0.9233593391463975,
best_f1=0.9255813953488372
```

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

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

```
step: 0, loss: 0.007147696800529957  
step: 10, loss: 0.01807512529194355  
step: 20, loss: 0.0023046929854899645  
step: 30, loss: 0.0011266516521573067  
step: 40, loss: 0.0035248370841145515  
step: 50, loss: 0.00786778423935175  
step: 60, loss: 0.0012444169260561466  
step: 70, loss: 0.00409657834097743  
step: 80, loss: 0.011403522454202175  
step: 90, loss: 0.10089017450809479  
step: 100, loss: 0.006852279417216778  
step: 110, loss: 0.008939215913414955  
step: 120, loss: 0.0008425600826740265  
step: 130, loss: 0.0012486231280490756  
step: 140, loss: 0.0068779680877923965  
step: 150, loss: 0.0008938902174122632  
step: 160, loss: 0.044819802045822144  
step: 170, loss: 0.03900863975286484  
step: 180, loss: 0.0020068297162652016  
step: 190, loss: 0.030976194888353348  
step: 200, loss: 0.0007508082780987024  
step: 210, loss: 0.029280373826622963  
step: 220, loss: 0.004640243016183376  
step: 230, loss: 0.01454614382237196  
step: 240, loss: 0.004658445715904236  
step: 250, loss: 0.0548381581902504  
step: 260, loss: 0.005705665796995163  
step: 270, loss: 0.004296662751585245  
step: 280, loss: 0.07448897510766983  
step: 290, loss: 0.01502649299800396  
step: 300, loss: 0.0006727893487550318  
step: 310, loss: 0.020743699744343758  
step: 320, loss: 0.02325860783457756  
step: 330, loss: 0.01478403341025114  
step: 340, loss: 0.026370733976364136  
step: 350, loss: 0.0025042248889803886  
step: 360, loss: 0.06618388742208481  
step: 370, loss: 0.007418269291520119  
step: 380, loss: 0.13208293914794922  
step: 390, loss: 0.0017776162130758166  
step: 400, loss: 0.004025349393486977  
step: 410, loss: 0.004856453742831945  
step: 420, loss: 0.0030742811504751444  
step: 430, loss: 0.07455247640609741  
step: 440, loss: 0.002155161462724209
```

```

step: 450, loss: 0.013979723677039146
step: 460, loss: 0.006869500502943993
step: 470, loss: 0.02723110094666481
step: 480, loss: 0.0033924258314073086
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 490, loss: 0.00847608782351017
step: 500, loss: 0.0011300548212602735
step: 510, loss: 0.09867612272500992
step: 520, loss: 0.09129323065280914
step: 530, loss: 0.02367904596030712
/usr/local/lib/python3.7/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=0.921747042766151, f1=0.9239819004524887,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00417288625612855
step: 10, loss: 0.05244938284158707
step: 20, loss: 0.002027304144576192
step: 30, loss: 0.0005158873391337693
step: 40, loss: 0.0022445598151534796
step: 50, loss: 0.004215126857161522
step: 60, loss: 0.010447892360389233
step: 70, loss: 0.0035119615495204926
step: 80, loss: 0.001543457736261189
step: 90, loss: 0.0056984093971550465
step: 100, loss: 0.0010288163321092725
step: 110, loss: 0.0006413617520593107
step: 120, loss: 0.0006378149264492095
step: 130, loss: 0.000354765506926924
step: 140, loss: 0.003047107020393014
step: 150, loss: 0.029712116345763206
step: 160, loss: 0.0013482411159202456
step: 170, loss: 0.020329104736447334
step: 180, loss: 0.0009253494790755212
step: 190, loss: 0.009276410564780235
step: 200, loss: 0.0003714756458066404
step: 210, loss: 0.005554028321057558
step: 220, loss: 0.002237962791696191
step: 230, loss: 0.007042642682790756
step: 240, loss: 0.001348467543721199
step: 250, loss: 0.014347410760819912
step: 260, loss: 0.13856473565101624
step: 270, loss: 0.004956948570907116
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 280, loss: 0.14128409326076508

```

```

step: 290, loss: 0.11031996458768845
step: 300, loss: 0.023156071081757545
step: 310, loss: 0.03822166845202446
step: 320, loss: 0.12638042867183685
step: 330, loss: 0.007724458817392588
step: 340, loss: 0.015851283445954323
step: 350, loss: 0.006770619656890631
step: 360, loss: 0.012443670071661472
step: 370, loss: 0.04467493295669556
step: 380, loss: 0.0005967324832454324
step: 390, loss: 0.0033597806468605995
step: 400, loss: 0.012258922681212425
step: 410, loss: 0.010162794962525368
step: 420, loss: 0.006574089638888836
step: 430, loss: 0.12355057895183563
step: 440, loss: 0.003159078536555171
step: 450, loss: 0.0049308594316244125
step: 460, loss: 0.005865391809493303
step: 470, loss: 0.006567738018929958
step: 480, loss: 0.0014432905009016395
step: 490, loss: 0.0006657705525867641
step: 500, loss: 0.006711837835609913
step: 510, loss: 0.009480474516749382
step: 520, loss: 0.0015568924136459827
step: 530, loss: 0.00045568670611828566
/usr/local/lib/python3.7/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=0.9139382600561271, f1=0.909689557855127,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015599660109728575
step: 10, loss: 0.0004653633222915232
step: 20, loss: 0.14449095726013184
step: 30, loss: 0.0013869327958673239
step: 40, loss: 0.20027923583984375
step: 50, loss: 0.015238335356116295
step: 60, loss: 0.00020227617642376572
step: 70, loss: 0.00032952739275060594
step: 80, loss: 0.00470007536932826
step: 90, loss: 0.00025759427808225155
step: 100, loss: 0.0006160447956062853
step: 110, loss: 0.0007862620404921472
step: 120, loss: 0.00016527553088963032
step: 130, loss: 0.02716093696653843
step: 140, loss: 0.08090932667255402

```

```

step: 150, loss: 0.03802022337913513
step: 160, loss: 0.0020993989892303944
step: 170, loss: 0.010723892599344254
step: 180, loss: 0.0007802923792041838
step: 190, loss: 0.0016676095547154546
step: 200, loss: 0.00048780511133372784
step: 210, loss: 0.00041056869667954743
step: 220, loss: 0.008394970558583736
step: 230, loss: 0.00020772541756741703
step: 240, loss: 0.0006851161597296596
step: 250, loss: 0.0003178355982527137
step: 260, loss: 8.112853538477793e-05
step: 270, loss: 0.08883346617221832
step: 280, loss: 0.003090959507972002
step: 290, loss: 0.000682882615365088
step: 300, loss: 0.003953041508793831
step: 310, loss: 0.0036102079320698977
step: 320, loss: 0.00017191485676448792
step: 330, loss: 0.0003412209334783256
step: 340, loss: 0.05522080883383751
step: 350, loss: 0.00244011334143579
step: 360, loss: 0.002627623500302434
step: 370, loss: 0.00033817681833170354
step: 380, loss: 0.00018318393267691135
step: 390, loss: 0.005219448357820511
step: 400, loss: 0.00014410474977921695
step: 410, loss: 0.0008545488817617297
step: 420, loss: 0.005240589380264282
step: 430, loss: 0.05005137249827385
step: 440, loss: 0.0003272552276030183
step: 450, loss: 0.06046310067176819
step: 460, loss: 0.00013414869317784905
step: 470, loss: 0.017489677295088768
step: 480, loss: 0.0836324542760849
step: 490, loss: 0.0161756444722414
step: 500, loss: 0.05687301978468895
step: 510, loss: 0.021256260573863983
step: 520, loss: 0.010260879062116146
step: 530, loss: 0.018933530896902084
/usr/local/lib/python3.7/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=0.9209654519640322, f1=0.9170222854433381,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003267022257205099

```


step: 10, loss: 0.014091124758124352
step: 20, loss: 0.0011409538565203547
step: 30, loss: 0.002920610597357154
step: 40, loss: 0.004994492512196302
step: 50, loss: 0.0011790733551606536
step: 60, loss: 0.0005727813113480806
step: 70, loss: 0.0010239561088383198
step: 80, loss: 0.17004770040512085
step: 90, loss: 0.0010157234501093626
step: 100, loss: 0.006787844933569431
step: 110, loss: 0.04781564697623253
step: 120, loss: 0.045903149992227554
step: 130, loss: 0.0013818478910252452
step: 140, loss: 0.00011832106247311458
step: 150, loss: 0.0004498297639656812
step: 160, loss: 0.00013712547661270946
step: 170, loss: 0.00044585950672626495
step: 180, loss: 0.0008168102358467877
step: 190, loss: 0.0033272448927164078
step: 200, loss: 0.0015867482870817184
step: 210, loss: 0.018228260800242424
step: 220, loss: 0.014602185226976871
step: 230, loss: 0.0009163021459244192
step: 240, loss: 0.006412621587514877
step: 250, loss: 0.0016625265125185251
step: 260, loss: 0.0007569349254481494
step: 270, loss: 0.0003732126497197896
step: 280, loss: 0.00020964507712051272
step: 290, loss: 0.017134172841906548
step: 300, loss: 0.0006988910026848316
step: 310, loss: 0.00014470475434791297
step: 320, loss: 0.0006041557062417269
step: 330, loss: 0.0001373492559650913
step: 340, loss: 0.061263371258974075
step: 350, loss: 0.012861969880759716
step: 360, loss: 0.0037551256828010082
step: 370, loss: 0.0014462519902735949
step: 380, loss: 0.008173920214176178
step: 390, loss: 0.0003855710965581238
step: 400, loss: 0.00730011984705925
step: 410, loss: 0.0011353137670084834
step: 420, loss: 0.013084765523672104
step: 430, loss: 0.00047709685168229043
step: 440, loss: 0.0662645474076271
step: 450, loss: 0.00030535657424479723
step: 460, loss: 0.0006776078953407705
step: 470, loss: 0.002018702682107687
step: 480, loss: 0.12060173600912094

```
step: 490, loss: 0.018439220264554024
step: 500, loss: 0.08258359879255295
step: 510, loss: 0.0007879474433138967
step: 520, loss: 0.08973533660173416
step: 530, loss: 0.010261191986501217
/usr/local/lib/python3.7/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=0.9202226345083488, f1=0.9229349330872173,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006693791947327554
step: 10, loss: 0.049602583050727844
step: 20, loss: 0.0003563402278814465
step: 30, loss: 0.002390509471297264
step: 40, loss: 0.0007470845011994243
step: 50, loss: 0.0015698939096182585
step: 60, loss: 0.0021868585608899593
step: 70, loss: 0.0002040662948274985
step: 80, loss: 0.01444171741604805
step: 90, loss: 0.0005674426211044192
step: 100, loss: 0.00031355838291347027
step: 110, loss: 0.001139483298175037
step: 120, loss: 0.00012704834807664156
step: 130, loss: 0.00019406976934988052
step: 140, loss: 0.005558306816965342
step: 150, loss: 0.0001094965118682012
step: 160, loss: 0.0002796879271045327
step: 170, loss: 0.00024063534510787576
step: 180, loss: 0.00012363225687295198
step: 190, loss: 0.002632085932418704
step: 200, loss: 0.0008551536593586206
step: 210, loss: 0.0001237105461768806
step: 220, loss: 0.002309110714122653
step: 230, loss: 0.0017704465426504612
step: 240, loss: 0.00022684779833070934
step: 250, loss: 0.0005361834773793817
step: 260, loss: 0.0002053577482001856
step: 270, loss: 0.003123078728094697
step: 280, loss: 0.028369473293423653
step: 290, loss: 0.0012992058182135224
step: 300, loss: 0.005419852212071419
step: 310, loss: 0.0014072256162762642
step: 320, loss: 0.0002243217604700476
step: 330, loss: 0.0019973127637058496
step: 340, loss: 0.00043511835974641144
```

```

step: 350, loss: 4.27947161369957e-05
step: 360, loss: 0.0064310417510569096
step: 370, loss: 0.00010441470658406615
step: 380, loss: 0.0008302011410705745
step: 390, loss: 0.10772240161895752
step: 400, loss: 0.003030004445463419
step: 410, loss: 4.807504956261255e-05
step: 420, loss: 0.10167519748210907
step: 430, loss: 0.03852329030632973
step: 440, loss: 0.00016827329818625003
step: 450, loss: 0.0010515687754377723
step: 460, loss: 0.00019102566875517368
step: 470, loss: 0.000353988230926916
step: 480, loss: 0.00031083147041499615
step: 490, loss: 0.004347864538431168
step: 500, loss: 0.02569190412759781
step: 510, loss: 0.039159372448921204
step: 520, loss: 0.0020289032254368067
step: 530, loss: 0.0007130238227546215
/usr/local/lib/python3.7/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=0.9199623352165726, f1=0.9218604651162791,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026119156973436475
step: 10, loss: 8.489147876389325e-05
step: 20, loss: 0.000989857711829245
step: 30, loss: 0.0021844012662768364
step: 40, loss: 6.114480493124574e-05
step: 50, loss: 0.0007658420945517719
step: 60, loss: 0.00045348572893999517
step: 70, loss: 0.0030114087276160717
step: 80, loss: 0.02440389059484005
step: 90, loss: 6.765813304809853e-05
step: 100, loss: 0.00014331124839372933
step: 110, loss: 0.0005200206651352346
step: 120, loss: 0.00013761578884441406
step: 130, loss: 0.0001260633289348334
step: 140, loss: 0.00013211513578426093
step: 150, loss: 6.116841541370377e-05
step: 160, loss: 0.0001500084763392806
step: 170, loss: 0.032153982669115067
step: 180, loss: 0.00042749595013447106
step: 190, loss: 9.199719352182001e-05
step: 200, loss: 0.05196329578757286

```

```

step: 210, loss: 0.09146367013454437
step: 220, loss: 0.003333291970193386
step: 230, loss: 0.00028914763242937624
step: 240, loss: 0.0004915977478958666
step: 250, loss: 0.07686765491962433
step: 260, loss: 0.000442476972239092
step: 270, loss: 0.00014984187146183103
step: 280, loss: 0.0001217009048559703
step: 290, loss: 0.051293887197971344
step: 300, loss: 9.289749868912622e-05
step: 310, loss: 0.012810231186449528
step: 320, loss: 0.0011418252252042294
step: 330, loss: 0.0013044063234701753
step: 340, loss: 0.0005322231445461512
step: 350, loss: 0.0034357605036348104
step: 360, loss: 4.875038575846702e-05
step: 370, loss: 0.005772207863628864
step: 380, loss: 0.01721269264817238
step: 390, loss: 0.014950133860111237
step: 400, loss: 0.0002895992947742343
step: 410, loss: 0.0003685754491016269
step: 420, loss: 5.243811756372452e-05
step: 430, loss: 0.0003645193064585328
step: 440, loss: 0.0016640170942991972
step: 450, loss: 0.08483187109231949
step: 460, loss: 0.00021654072043020278
step: 470, loss: 3.7247755244607106e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 480, loss: 0.00018905494653154165
step: 490, loss: 0.0036559998989105225
step: 500, loss: 0.0008387192501686513
step: 510, loss: 0.006442230194807053
step: 520, loss: 0.0002443782286718488
step: 530, loss: 0.00047863213694654405
/usr/local/lib/python3.7/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=0.922227335480902, f1=0.9221556886227545,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009462616872042418
step: 10, loss: 0.00019566611445043236
step: 20, loss: 8.642298053018749e-05
step: 30, loss: 9.79123724391684e-05
step: 40, loss: 0.0007222218555398285
step: 50, loss: 8.039467502385378e-05

```

step: 60, loss: 0.0003479671140667051
step: 70, loss: 0.0029582465067505836
step: 80, loss: 3.7050289392936975e-05
step: 90, loss: 5.1482820708770305e-05
step: 100, loss: 4.795621134690009e-05
step: 110, loss: 0.00011032228212570772
step: 120, loss: 7.149628072511405e-05
step: 130, loss: 0.00037073102430440485
step: 140, loss: 0.0027653079014271498
step: 150, loss: 0.0001610834733583033
step: 160, loss: 0.00011535235535120592
step: 170, loss: 0.0038459303323179483
step: 180, loss: 0.0001642486749915406
step: 190, loss: 0.00023275452258531004
step: 200, loss: 0.002240004250779748
step: 210, loss: 0.0072290729731321335
step: 220, loss: 0.0008126195752993226
step: 230, loss: 0.05791519209742546
step: 240, loss: 7.482572982553393e-05
step: 250, loss: 0.00018234702292829752
step: 260, loss: 0.010573787614703178
step: 270, loss: 0.00040343982982449234
step: 280, loss: 0.0014186688931658864
step: 290, loss: 0.0001655960368225351
step: 300, loss: 4.7407047532033175e-05
step: 310, loss: 0.00027773340116254985
step: 320, loss: 0.002476580673828721
step: 330, loss: 0.0025478580500930548
step: 340, loss: 0.0006555697182193398
step: 350, loss: 0.00018750583694782108
step: 360, loss: 0.00010164533159695566
step: 370, loss: 0.0044004470109939575
step: 380, loss: 0.0013471177080646157
step: 390, loss: 0.0002517283137422055
step: 400, loss: 0.001586145255714655
step: 410, loss: 0.0030226819217205048
step: 420, loss: 0.0004669151676353067
step: 430, loss: 2.6448939024703577e-05
step: 440, loss: 0.00025886762887239456
step: 450, loss: 0.0010763118043541908
step: 460, loss: 0.000404568447265774
step: 470, loss: 0.0005303098587319255
step: 480, loss: 3.577999450499192e-05
step: 490, loss: 0.0008593475795350969
step: 500, loss: 9.265223343390971e-05
step: 510, loss: 3.7488480302272364e-05
step: 520, loss: 7.31639884179458e-05
step: 530, loss: 7.435619045281783e-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=0.9230046948356808, f1=0.9231485794131347,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.859556090901606e-05
step: 10, loss: 0.00022324993915390223
step: 20, loss: 0.00016757268167566508
step: 30, loss: 0.00020788579422514886
step: 40, loss: 8.535103552276269e-05
step: 50, loss: 0.007973022758960724
step: 60, loss: 0.004668578505516052
step: 70, loss: 3.541441401466727e-05
step: 80, loss: 3.6230398109182715e-05
step: 90, loss: 5.545494059333578e-05
step: 100, loss: 0.0003399615525268018
step: 110, loss: 0.00015955873823259026
step: 120, loss: 2.937291719717905e-05
step: 130, loss: 4.35886686318554e-05
step: 140, loss: 0.00036274309968575835
step: 150, loss: 2.60311717283912e-05
step: 160, loss: 8.933877688832581e-05
step: 170, loss: 0.017418503761291504
step: 180, loss: 8.819871436571702e-05
step: 190, loss: 0.02069076895713806
step: 200, loss: 0.00344204343855381
step: 210, loss: 0.0006994273862801492
step: 220, loss: 0.0008979755220934749
step: 230, loss: 2.189290535170585e-05
step: 240, loss: 0.00020057277288287878
step: 250, loss: 3.602532160584815e-05
step: 260, loss: 3.0385839636437595e-05
step: 270, loss: 5.1802388043142855e-05
step: 280, loss: 5.630352097796276e-05
step: 290, loss: 6.170627602841705e-05
step: 300, loss: 9.82040073722601e-05
step: 310, loss: 0.00012652372242882848
step: 320, loss: 4.890144555247389e-05
step: 330, loss: 5.512115967576392e-05
step: 340, loss: 3.593735891627148e-05
step: 350, loss: 0.0004967285203747451
step: 360, loss: 1.9531311409082264e-05
step: 370, loss: 0.008911370299756527
step: 380, loss: 5.391496961237863e-05
step: 390, loss: 2.562506779213436e-05

```

```

step: 400, loss: 0.0005644229822792113
step: 410, loss: 0.0001961792731890455
step: 420, loss: 3.202167135896161e-05
step: 430, loss: 3.249401561333798e-05
step: 440, loss: 4.776802597916685e-05
step: 450, loss: 2.966332613141276e-05
step: 460, loss: 0.07672582566738129
step: 470, loss: 2.8777129045920447e-05
step: 480, loss: 0.001860781223513186
step: 490, loss: 2.8787912015104666e-05
step: 500, loss: 0.00040003270260058343
step: 510, loss: 0.00017792989092413336
step: 520, loss: 0.00018432934302836657
step: 530, loss: 7.490058487746865e-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=0.9209431345353675, f1=0.9230064161319891,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000158315320732072
step: 10, loss: 1.9374932890059426e-05
step: 20, loss: 3.0475332096102647e-05
step: 30, loss: 0.000592017313465476
step: 40, loss: 4.6812983782729134e-05
step: 50, loss: 0.01905723288655281
step: 60, loss: 0.0008031734032556415
step: 70, loss: 4.0249356970889494e-05
step: 80, loss: 3.2326061045750976e-05
step: 90, loss: 5.932448402745649e-05
step: 100, loss: 0.00045740348286926746
step: 110, loss: 3.2475643820362166e-05
step: 120, loss: 1.9635617718449794e-05
step: 130, loss: 4.299716238165274e-05
step: 140, loss: 5.428283111541532e-05
step: 150, loss: 1.8190283299190924e-05
step: 160, loss: 1.837282434280496e-05
step: 170, loss: 2.5498731702100486e-05
step: 180, loss: 2.5070425181183964e-05
step: 190, loss: 0.002298591425642371
step: 200, loss: 4.8309186240658164e-05
step: 210, loss: 0.00035746931098401546
step: 220, loss: 0.00038766604848206043
step: 230, loss: 1.7288786693825386e-05
step: 240, loss: 0.00010036714957095683
step: 250, loss: 2.166950434911996e-05

```

```

step: 260, loss: 2.0708439478767104e-05
step: 270, loss: 3.7802554288646206e-05
step: 280, loss: 0.0038034708704799414
step: 290, loss: 2.462725569785107e-05
step: 300, loss: 3.087427467107773e-05
step: 310, loss: 0.0006537591107189655
step: 320, loss: 0.08988799154758453
step: 330, loss: 4.6789104089839384e-05
step: 340, loss: 0.0013822474284097552
step: 350, loss: 9.344496356789023e-05
step: 360, loss: 9.193099685944617e-05
step: 370, loss: 0.005769773852080107
step: 380, loss: 0.0004318663850426674
step: 390, loss: 0.00010037099127657712
step: 400, loss: 4.720276047009975e-05
step: 410, loss: 0.00029901223024353385
step: 420, loss: 0.0009351748158223927
step: 430, loss: 0.0014861890813335776
step: 440, loss: 2.0157180188107304e-05
step: 450, loss: 0.0009295609779655933
step: 460, loss: 2.310371201019734e-05
step: 470, loss: 0.0001690041390247643
step: 480, loss: 0.00022073349100537598
step: 490, loss: 0.00810534693300724
step: 500, loss: 0.00012321150279603899
step: 510, loss: 5.3636173106497154e-05
step: 520, loss: 2.5510136765660718e-05
step: 530, loss: 2.838958243955858e-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=0.9237209302325581, f1=0.9254143646408841,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001498408819315955
step: 10, loss: 4.3260337406536564e-05
step: 20, loss: 2.5441313482588157e-05
step: 30, loss: 9.449812932871282e-05
step: 40, loss: 0.0002109076303895563
step: 50, loss: 8.432959293713793e-05
step: 60, loss: 0.00012723403051495552
step: 70, loss: 6.013601887389086e-05
step: 80, loss: 0.00044530763989314437
step: 90, loss: 0.03053693100810051
step: 100, loss: 0.0001237498945556581
step: 110, loss: 3.357769310241565e-05

```



```

step: 120, loss: 5.870494351256639e-05
step: 130, loss: 3.175426536472514e-05
step: 140, loss: 2.3937929654493928e-05
step: 150, loss: 5.420799061539583e-05
step: 160, loss: 1.6800819139461964e-05
step: 170, loss: 0.00014190946239978075
step: 180, loss: 4.1410752601223066e-05
step: 190, loss: 6.20688297203742e-05
step: 200, loss: 1.8596341760712676e-05
step: 210, loss: 2.7158817829331383e-05
step: 220, loss: 1.613763379282318e-05
step: 230, loss: 2.206767203460913e-05
step: 240, loss: 2.79125124507118e-05
step: 250, loss: 1.5832263670745306e-05
step: 260, loss: 1.673360566201154e-05
step: 270, loss: 1.9132503439323045e-05
step: 280, loss: 2.3908329239930026e-05
step: 290, loss: 0.00012493114627432078
step: 300, loss: 3.6740351788466796e-05
step: 310, loss: 2.872448203561362e-05
step: 320, loss: 0.0006515937857329845
step: 330, loss: 4.142413308727555e-05
step: 340, loss: 0.00010598474182188511
step: 350, loss: 1.6040905393310823e-05
step: 360, loss: 2.847773430403322e-05
step: 370, loss: 3.0482518923236057e-05
step: 380, loss: 1.601854819455184e-05
step: 390, loss: 0.00010451948764966801
step: 400, loss: 3.0119306757114828e-05
step: 410, loss: 0.00013617232616525143
step: 420, loss: 2.0403040252858773e-05
step: 430, loss: 1.7843856767285615e-05
step: 440, loss: 1.8782515326165594e-05
step: 450, loss: 2.0261522877262905e-05
step: 460, loss: 6.029401993146166e-05
step: 470, loss: 0.00017784589726943523
step: 480, loss: 5.162858360563405e-05
step: 490, loss: 2.1241232389002107e-05
step: 500, loss: 1.6629393940092996e-05
step: 510, loss: 1.8871849533752538e-05
step: 520, loss: 0.008398296311497688
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 530, loss: 1.5329369489336386e-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=0.9196386115073705, f1=0.9226441631504922,
best_f1=0.9255813953488372

```

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

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

step: 0, loss: 6.904284236952662e-05
step: 10, loss: 2.2417209038394503e-05
step: 20, loss: 1.5925386833259836e-05
step: 30, loss: 1.760543636919465e-05
step: 40, loss: 2.8906873922096565e-05
step: 50, loss: 0.0009055513655766845
step: 60, loss: 1.3172501894587185e-05
step: 70, loss: 0.00010071021824842319
step: 80, loss: 1.8052458472084254e-05
step: 90, loss: 0.0027154714334756136
step: 100, loss: 2.948472865682561e-05
step: 110, loss: 0.0006972627597860992
step: 120, loss: 4.54734472441487e-05
step: 130, loss: 0.002242202404886484
step: 140, loss: 2.3039363441057503e-05
step: 150, loss: 0.04795358330011368
step: 160, loss: 8.346047980012372e-05
step: 170, loss: 0.00010517404007259756
step: 180, loss: 7.690843631280586e-05
step: 190, loss: 1.8521825040807016e-05
step: 200, loss: 0.0001232760405400768
step: 210, loss: 2.1043826563982293e-05
step: 220, loss: 3.55545780621469e-05
step: 230, loss: 4.5675515139009804e-05
step: 240, loss: 1.599619645276107e-05
step: 250, loss: 3.807723260251805e-05
step: 260, loss: 6.490547093562782e-05
step: 270, loss: 2.035071338468697e-05
step: 280, loss: 1.6145173503900878e-05
step: 290, loss: 0.006077332887798548
step: 300, loss: 2.6072269974974915e-05
step: 310, loss: 3.415410537854768e-05
step: 320, loss: 3.227424167562276e-05
step: 330, loss: 3.553950955392793e-05
step: 340, loss: 0.0012664432870224118
step: 350, loss: 7.713820377830416e-05
step: 360, loss: 0.00016718890401534736
step: 370, loss: 3.0079665521043353e-05
step: 380, loss: 1.832070847740397e-05
step: 390, loss: 0.00014718499733135104
step: 400, loss: 2.312208744115196e-05
step: 410, loss: 2.047715497610625e-05
step: 420, loss: 2.2343343516695313e-05
step: 430, loss: 2.581149965408258e-05
step: 440, loss: 0.0006853666855022311

```

step: 450, loss: 2.5103934603976086e-05
step: 460, loss: 2.6831527065951377e-05
step: 470, loss: 0.00011836967314593494
step: 480, loss: 2.7029242119169794e-05
step: 490, loss: 1.8778786397888325e-05
step: 500, loss: 0.00010351071978220716
step: 510, loss: 0.0005516072269529104
step: 520, loss: 1.7825264876591973e-05
step: 530, loss: 3.098437809967436e-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=0.9201520912547528, f1=0.922138836772983,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.6294181477860548e-05
step: 10, loss: 0.0055277710780501366
step: 20, loss: 1.5761475879116915e-05
step: 30, loss: 1.9043196516577154e-05
step: 40, loss: 0.03234041854739189
step: 50, loss: 3.492823088890873e-05
step: 60, loss: 3.419392669457011e-05
step: 70, loss: 1.7515856598038226e-05
step: 80, loss: 3.458589344518259e-05
step: 90, loss: 2.9830316634615883e-05
step: 100, loss: 7.279696001205593e-05
step: 110, loss: 0.00016756770492065698
step: 120, loss: 2.4577513613621704e-05
step: 130, loss: 0.026398295536637306
step: 140, loss: 1.8793574781739153e-05
step: 150, loss: 2.4671673600096256e-05
step: 160, loss: 0.00023425866675097495
step: 170, loss: 1.5109590094652958e-05
step: 180, loss: 1.5135675312194508e-05
step: 190, loss: 3.708943768288009e-05
step: 200, loss: 2.27862528845435e-05
step: 210, loss: 2.1118128643138334e-05
step: 220, loss: 0.00016250890621449798
step: 230, loss: 7.386923243757337e-05
step: 240, loss: 1.6718870028853416e-05
step: 250, loss: 4.8293863073922694e-05
step: 260, loss: 2.0272607798688114e-05
step: 270, loss: 1.776184763002675e-05
step: 280, loss: 2.38075135712279e-05
step: 290, loss: 1.2062382666044869e-05
step: 300, loss: 0.0001778291625669226

```

```

step: 310, loss: 0.0005580346332862973
step: 320, loss: 5.957342727924697e-05
step: 330, loss: 5.77381688344758e-05
step: 340, loss: 0.023778146132826805
step: 350, loss: 1.2095915735699236e-05
step: 360, loss: 0.00010670811025192961
step: 370, loss: 1.2107066140742972e-05
step: 380, loss: 2.6824338419828564e-05
step: 390, loss: 2.352824594709091e-05
step: 400, loss: 3.467773422016762e-05
step: 410, loss: 6.072026371839456e-05
step: 420, loss: 1.8752736650640145e-05
step: 430, loss: 3.5748515074374154e-05
step: 440, loss: 7.581592944916338e-05
step: 450, loss: 1.478917238273425e-05
step: 460, loss: 1.661080023040995e-05
step: 470, loss: 0.005726093892008066
step: 480, loss: 1.2017682820442133e-05
step: 490, loss: 1.359344878437696e-05
step: 500, loss: 1.7895979908644222e-05
step: 510, loss: 4.00702738261316e-05
step: 520, loss: 1.7746813682606444e-05
step: 530, loss: 1.6823196347104385e-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=0.9216241737488197, f1=0.9227188081936686,
best_f1=0.9255813953488372
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 [00:26, 218.74it/s]
load_f1 = 0.9264504339881224
real_f1 = 0.9245454545454544
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 237.57it/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 446 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: 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: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: 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: 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: 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: 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: 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: 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: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: plac<1.2.0,>=0.9.6 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.1.3)
Requirement already satisfied: 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: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg

```

```

Building wheel for en-core-web-lg (setup.py) ... done
Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-any.whl size=829180942
sha256=ac8c5d709e640e17e6964f3d0c0ecc7da918ba80d38e017805ee190794279aff
Stored in directory: /tmp/pip-ephem-wheel-cache-dmhn7lmh/wheels/11/95/ba/2c36cc368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

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

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

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

```

Selected optimization level 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: 0.5399580597877502

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

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

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

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=0.39999999999999997, f1=0.3902439024390244,

best_f1=0.3902439024390244

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


```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5181444883346558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4210526315789473, f1=0.4210526315789473,
best_f1=0.4210526315789473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48835495114326477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5333333333333333, f1=0.47058823529411764,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2409723848104477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6153846153846153, f1=0.4827586206896552,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2597409784793854
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.4545454545454545,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2313908040523529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7096774193548386, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3094281852245331
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5833333333333334, 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: 0.022965753450989723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8148148148148148, f1=0.6086956521739131,
best_f1=0.6086956521739131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007281591650098562
/usr/local/lib/python3.7/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=0.7741935483870968, f1=0.5925925925925927,
best_f1=0.6086956521739131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.04734592139720917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8461538461538461, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005917887669056654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8275862068965518, f1=0.6923076923076924,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007738039828836918
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.7200000000000001,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024002982303500175
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004319198429584503
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00759706599637866
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.6666666666666666,
best_f1=0.6666666666666666
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, 120594.52it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.75
real_f1 = 0.75
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 251.38it/s]

```

3.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6063847541809082
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.6061411499977112
step: 20, loss: 0.3685799837112427
step: 30, loss: 0.15220880508422852
step: 40, loss: 0.270976185798645
step: 50, loss: 0.02663455903530121
step: 60, loss: 0.08278117328882217
step: 70, loss: 0.015009061433374882
step: 80, loss: 0.17317406833171844
step: 90, loss: 0.13275843858718872
step: 100, loss: 0.012678215280175209
step: 110, loss: 0.12765875458717346
step: 120, loss: 0.013671454973518848
step: 130, loss: 0.005463265813887119
step: 140, loss: 0.0059684268198907375
step: 150, loss: 0.030092477798461914
step: 160, loss: 0.00609867786988616
step: 170, loss: 0.02829534374177456
step: 180, loss: 0.1434885412454605
step: 190, loss: 0.010669614188373089
step: 200, loss: 0.025032879784703255
step: 210, loss: 0.001112479716539383
step: 220, loss: 0.0021224089432507753
step: 230, loss: 0.004328756593167782

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9785794813979707, f1=0.9773755656108598,
best_f1=0.9773755656108598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001899603521451354
step: 10, loss: 0.0020250759553164244
step: 20, loss: 0.21143968403339386
step: 30, loss: 0.15617091953754425
step: 40, loss: 0.029641656205058098
step: 50, loss: 0.005041164346039295
step: 60, loss: 0.0028664011042565107
step: 70, loss: 0.10056362301111221
step: 80, loss: 0.0018926889169961214
step: 90, loss: 0.0016596426721662283
step: 100, loss: 0.007318643853068352
step: 110, loss: 0.09462811797857285
step: 120, loss: 0.0028422879986464977
step: 130, loss: 0.0748332291841507
step: 140, loss: 0.0029331042896956205
step: 150, loss: 0.12167497724294662
step: 160, loss: 0.028181811794638634
step: 170, loss: 0.0010317668784409761
step: 180, loss: 0.005029001738876104
step: 190, loss: 0.012249182909727097
step: 200, loss: 0.002897469326853752
step: 210, loss: 0.0029828930273652077
step: 220, loss: 0.1259211152791977
step: 230, loss: 0.020810462534427643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830124575311437, f1=0.9807037457434733,
best_f1=0.9807037457434733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00666987057775259
step: 10, loss: 0.0026477924548089504
step: 20, loss: 0.1164061576128006
step: 30, loss: 0.046330858021974564
step: 40, loss: 0.021001433953642845
step: 50, loss: 0.012219641357660294
step: 60, loss: 0.006344345398247242
step: 70, loss: 0.004927016794681549
step: 80, loss: 0.0018767720321193337
step: 90, loss: 0.029960574582219124
step: 100, loss: 0.0007174228085204959
step: 110, loss: 0.010307785123586655
step: 120, loss: 0.013893971219658852
step: 130, loss: 0.0006371315103024244
step: 140, loss: 0.021411990746855736
step: 150, loss: 0.001640992471948266
step: 160, loss: 0.07829500734806061
step: 170, loss: 0.005387237761169672
step: 180, loss: 0.09708990156650543
step: 190, loss: 0.001917945221066475
step: 200, loss: 0.0031243334524333477
step: 210, loss: 0.02285945974290371
step: 220, loss: 0.0003163541841786355
step: 230, loss: 0.006429762579500675
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876265466816648, f1=0.9785794813979707,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007956188055686653
step: 10, loss: 0.00028051063418388367
step: 20, loss: 0.001177024794742465
step: 30, loss: 0.0005251789116300642
step: 40, loss: 0.0023219904396682978
step: 50, loss: 0.005127921234816313
step: 60, loss: 0.0006297455402091146
step: 70, loss: 0.00036437169183045626
step: 80, loss: 0.0069387792609632015
step: 90, loss: 0.022844715043902397
step: 100, loss: 0.000246578041696921
step: 110, loss: 0.0002935861411970109

```



```

step: 120, loss: 0.019844019785523415
step: 130, loss: 0.0005109617486596107
step: 140, loss: 0.002224996453151107
step: 150, loss: 0.11532366275787354
step: 160, loss: 0.003701903158798814
step: 170, loss: 0.03359363228082657
step: 180, loss: 0.0004123346880078316
step: 190, loss: 0.002392654772847891
step: 200, loss: 0.0008382691303268075
step: 210, loss: 0.09556544572114944
step: 220, loss: 0.0008332289871759713
step: 230, loss: 0.002555268816649914
/usr/local/lib/python3.7/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=0.9798206278026906, f1=0.9818594104308391,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018815144896507263
step: 10, loss: 0.0004868492833338678
step: 20, loss: 0.0004923351225443184
step: 30, loss: 0.00019559079373721033
step: 40, loss: 0.00031057168962433934
step: 50, loss: 0.001521014142781496
step: 60, loss: 0.0003275475464761257
step: 70, loss: 0.0003708798612933606
step: 80, loss: 0.0002884118293877691
step: 90, loss: 0.00030010921182110906
step: 100, loss: 0.00034975825110450387
step: 110, loss: 0.0007224296568892896
step: 120, loss: 8.966137829702348e-05
step: 130, loss: 0.0005885600112378597
step: 140, loss: 0.0019958140328526497
step: 150, loss: 0.00017578902770765126
step: 160, loss: 0.00019147383864037693
step: 170, loss: 0.0006681001977995038
step: 180, loss: 0.007808004971593618
step: 190, loss: 0.0010558319045230746
step: 200, loss: 0.0031172712333500385
step: 210, loss: 0.0007309294887818396
step: 220, loss: 0.002990997163578868
step: 230, loss: 0.012345770373940468
/usr/local/lib/python3.7/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=0.9852440408626559, f1=0.9737742303306728,

```

```

best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006998075405135751
step: 10, loss: 0.0005881024990230799
step: 20, loss: 0.0007286136387847364
step: 30, loss: 0.0018955371342599392
step: 40, loss: 0.0006243845564313233
step: 50, loss: 0.00021112922695465386
step: 60, loss: 0.00015527076902799308
step: 70, loss: 0.0006597628816962242
step: 80, loss: 0.0007787839858792722
step: 90, loss: 0.0007487722323276103
step: 100, loss: 0.00420162221416831
step: 110, loss: 0.0003251529997214675
step: 120, loss: 0.0005905637517571449
step: 130, loss: 0.0010126373963430524
step: 140, loss: 0.00013550466974265873
step: 150, loss: 0.00028677110094577074
step: 160, loss: 0.00013966215192340314
step: 170, loss: 0.0003663020033854991
step: 180, loss: 0.023475488647818565
step: 190, loss: 0.0007809224771335721
step: 200, loss: 0.00029647519113495946
step: 210, loss: 0.00018461240688338876
step: 220, loss: 9.850153583101928e-05
step: 230, loss: 0.000474083935841918
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9785310734463276,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009023254737257957
step: 10, loss: 0.00013139276416040957
step: 20, loss: 0.001262559206224978
step: 30, loss: 0.00967410858720541
step: 40, loss: 0.0005872877663932741
step: 50, loss: 0.0007532053277827799
step: 60, loss: 0.0004973275354132056
step: 70, loss: 0.02759391814470291
step: 80, loss: 0.0011397399939596653
step: 90, loss: 7.312034722417593e-05
step: 100, loss: 0.00011980327690253034
step: 110, loss: 0.00020830276480410248

```

```

step: 120, loss: 0.0001005638187052682
step: 130, loss: 7.764172914903611e-05
step: 140, loss: 0.00016838278679642826
step: 150, loss: 8.951219933805987e-05
step: 160, loss: 0.0164212454110384
step: 170, loss: 0.00010334682883694768
step: 180, loss: 9.8455224360805e-05
step: 190, loss: 8.727045496925712e-05
step: 200, loss: 0.022423136979341507
step: 210, loss: 9.359693649457768e-05
step: 220, loss: 0.0059393011033535
step: 230, loss: 9.968822269001976e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9806598407281, f1=0.9690011481056257,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.607373845530674e-05
step: 10, loss: 0.0001057421977748163
step: 20, loss: 6.580573972314596e-05
step: 30, loss: 0.0015777223743498325
step: 40, loss: 0.00026697112480178475
step: 50, loss: 0.00042082398431375623
step: 60, loss: 0.00014421824016608298
step: 70, loss: 0.009210532531142235
step: 80, loss: 0.001924631418660283
step: 90, loss: 0.0008942732820287347
step: 100, loss: 0.0005547929904423654
step: 110, loss: 0.001042295596562326
step: 120, loss: 0.004888433031737804
step: 130, loss: 0.00011044398706872016
step: 140, loss: 0.00015236614854075015
step: 150, loss: 0.0002942817227449268
step: 160, loss: 0.013323637656867504
step: 170, loss: 8.052949124248698e-05
step: 180, loss: 0.00010069934069178998
step: 190, loss: 0.00013287996989674866
step: 200, loss: 0.00010765706974780187
step: 210, loss: 0.00011236949649173766
step: 220, loss: 0.010166860185563564
step: 230, loss: 0.00011010862363036722
/usr/local/lib/python3.7/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=0.9832026875699889, f1=0.9730337078651685,

```

```

best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001792027847841382
step: 10, loss: 0.0004768144863191992
step: 20, loss: 0.001456386991776526
step: 30, loss: 0.000456984736956656
step: 40, loss: 0.004054442513734102
step: 50, loss: 7.883958460297436e-05
step: 60, loss: 0.002201489871367812
step: 70, loss: 0.0002919521648436785
step: 80, loss: 0.0001348484365735203
step: 90, loss: 0.00018775812350213528
step: 100, loss: 0.18967746198177338
step: 110, loss: 0.00013241283886600286
step: 120, loss: 5.352906009647995e-05
step: 130, loss: 0.000103837825008668
step: 140, loss: 0.00013265978486742824
step: 150, loss: 0.002388936933130026
step: 160, loss: 8.688506204634905e-05
step: 170, loss: 0.00020003458485007286
step: 180, loss: 0.0004674613301176578
step: 190, loss: 0.0001881149655673653
step: 200, loss: 0.00012398569379001856
step: 210, loss: 0.0008670422248542309
step: 220, loss: 0.00011553884542081505
step: 230, loss: 0.00044765754137188196
/usr/local/lib/python3.7/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=0.9749430523917996, f1=0.9691428571428571,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008946314919739962
step: 10, loss: 3.432746234466322e-05
step: 20, loss: 0.0001238096010638401
step: 30, loss: 7.901627395767719e-05
step: 40, loss: 0.0001064474563463591
step: 50, loss: 0.0005407283315435052
step: 60, loss: 0.0001196784942294471
step: 70, loss: 0.0003014247049577534
step: 80, loss: 7.86533928476274e-05
step: 90, loss: 0.00012215848255436867
step: 100, loss: 6.471421511378139e-05
step: 110, loss: 7.738789281575009e-05

```

```

step: 120, loss: 6.039192885509692e-05
step: 130, loss: 5.8486552006797865e-05
step: 140, loss: 0.0003396380925551057
step: 150, loss: 0.004827779717743397
step: 160, loss: 4.486887337407097e-05
step: 170, loss: 5.639555456582457e-05
step: 180, loss: 0.00021587792434729636
step: 190, loss: 0.00032612867653369904
step: 200, loss: 0.0001377348235109821
step: 210, loss: 6.536805449286476e-05
step: 220, loss: 0.00017361968639306724
step: 230, loss: 0.0005775373429059982
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9727891156462585,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013698027760256082
step: 10, loss: 0.00015703821554780006
step: 20, loss: 9.283278632210568e-05
step: 30, loss: 0.02472231164574623
step: 40, loss: 6.895277329022065e-05
step: 50, loss: 8.71946249390021e-05
step: 60, loss: 6.902339373482391e-05
step: 70, loss: 8.578492270316929e-05
step: 80, loss: 6.620824569836259e-05
step: 90, loss: 6.074267366784625e-05
step: 100, loss: 0.00010506133548915386
step: 110, loss: 0.00018982011533807963
step: 120, loss: 0.0009854203090071678
step: 130, loss: 3.814095543930307e-05
step: 140, loss: 4.832183549297042e-05
step: 150, loss: 0.002800141926854849
step: 160, loss: 5.8154138969257474e-05
step: 170, loss: 0.0010881945490837097
step: 180, loss: 7.326491322601214e-05
step: 190, loss: 4.9949790991377085e-05
step: 200, loss: 0.018946044147014618
step: 210, loss: 4.6638142521260306e-05
step: 220, loss: 7.925438694655895e-05
step: 230, loss: 6.434498209273443e-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=0.9863945578231292, f1=0.9692132269099202,

```

```

best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017349018889945
step: 10, loss: 6.0225327615626156e-05
step: 20, loss: 7.409753743559122e-05
step: 30, loss: 0.00030431844061240554
step: 40, loss: 8.31218931125477e-05
step: 50, loss: 0.00020908683654852211
step: 60, loss: 0.0001400218898197636
step: 70, loss: 5.86696551181376e-05
step: 80, loss: 0.00011583740706555545
step: 90, loss: 5.278277603792958e-05
step: 100, loss: 4.3762171117123216e-05
step: 110, loss: 9.095244604395702e-05
step: 120, loss: 7.559364166809246e-05
step: 130, loss: 5.7183751778211445e-05
step: 140, loss: 0.0004603086563292891
step: 150, loss: 6.638812919845805e-05
step: 160, loss: 7.183884736150503e-05
step: 170, loss: 6.441536970669404e-05
step: 180, loss: 0.0023438180796802044
step: 190, loss: 6.408223271137103e-05
step: 200, loss: 3.7735193473054096e-05
step: 210, loss: 0.002410706365481019
step: 220, loss: 7.426502997986972e-05
step: 230, loss: 0.0003553454880602658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9841986455981941, f1=0.9751131221719457,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003180274798069149
step: 10, loss: 8.369646093342453e-05
step: 20, loss: 8.602045272709802e-05
step: 30, loss: 0.00011296874436084181
step: 40, loss: 4.9889073125086725e-05
step: 50, loss: 8.503634307999164e-05
step: 60, loss: 5.449318268802017e-05
step: 70, loss: 0.0006650199647992849
step: 80, loss: 4.149390952079557e-05
step: 90, loss: 0.00015226256800815463
step: 100, loss: 2.8065400329069234e-05
step: 110, loss: 0.00014038642984814942

```

```

step: 120, loss: 0.0013773278333246708
step: 130, loss: 6.0866190324304625e-05
step: 140, loss: 6.257022323552519e-05
step: 150, loss: 6.313817721093073e-05
step: 160, loss: 3.52284187101759e-05
step: 170, loss: 5.431820318335667e-05
step: 180, loss: 4.758374780067243e-05
step: 190, loss: 3.702003232319839e-05
step: 200, loss: 0.000744032091461122
step: 210, loss: 4.431540219229646e-05
step: 220, loss: 3.751571421162225e-05
step: 230, loss: 2.7946392947342247e-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=0.9841986455981941, f1=0.9717514124293786,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.816734897554852e-05
step: 10, loss: 0.0003850820066872984
step: 20, loss: 7.19949312042445e-05
step: 30, loss: 5.4906548029975966e-05
step: 40, loss: 0.016278760507702827
step: 50, loss: 3.817903052549809e-05
step: 60, loss: 0.00020715651044156402
step: 70, loss: 9.233618038706481e-05
step: 80, loss: 6.46230619167909e-05
step: 90, loss: 5.054446592112072e-05
step: 100, loss: 4.545517367660068e-05
step: 110, loss: 6.83152029523626e-05
step: 120, loss: 2.9867576813558117e-05
step: 130, loss: 3.789222682826221e-05
step: 140, loss: 6.0304035287117586e-05
step: 150, loss: 3.5299515730002895e-05
step: 160, loss: 4.2878611566266045e-05
step: 170, loss: 2.972701986436732e-05
step: 180, loss: 3.8502646930282936e-05
step: 190, loss: 5.15921892656479e-05
step: 200, loss: 2.8866254069725983e-05
step: 210, loss: 5.7339289924129844e-05
step: 220, loss: 2.615824450913351e-05
step: 230, loss: 6.868191849207506e-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=0.9853107344632768, f1=0.971687429218573,

```

```

best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.740714964806102e-05
step: 10, loss: 2.4575143470428884e-05
step: 20, loss: 6.982921331655234e-05
step: 30, loss: 6.053338438505307e-05
step: 40, loss: 6.98641815688461e-05
step: 50, loss: 0.0007875044248066843
step: 60, loss: 4.7445962991332635e-05
step: 70, loss: 3.3217278541997075e-05
step: 80, loss: 0.00025108145200647414
step: 90, loss: 9.456834231968969e-05
step: 100, loss: 3.496420322335325e-05
step: 110, loss: 3.401039430173114e-05
step: 120, loss: 5.795904144179076e-05
step: 130, loss: 0.00014509986795019358
step: 140, loss: 3.437956183915958e-05
step: 150, loss: 7.134519546525553e-05
step: 160, loss: 0.001298868446610868
step: 170, loss: 3.261746678617783e-05
step: 180, loss: 0.00021593444398604333
step: 190, loss: 9.861744911177084e-05
step: 200, loss: 0.00019570121366996318
step: 210, loss: 3.690890298457816e-05
step: 220, loss: 4.427408930496313e-05
step: 230, loss: 4.0987750253407285e-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=0.9841986455981941, f1=0.971687429218573,
best_f1=0.9785794813979707
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:13, 184.90it/s]
load_f1 = 0.9898762654668166
real_f1 = 0.9854096520763187
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 232.93it/s]

```

3.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

step: 0, loss: 0.5927411317825317
step: 10, loss: 0.5169671773910522
step: 20, loss: 0.5379564762115479
step: 30, loss: 0.1341688483953476
step: 40, loss: 0.1388813853263855
step: 50, loss: 0.2583771347999573
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.10012322664260864
step: 70, loss: 0.11278549581766129
step: 80, loss: 0.03795517981052399
step: 90, loss: 0.8866406083106995
step: 100, loss: 0.09932714700698853
step: 110, loss: 0.07061970233917236
step: 120, loss: 0.12735070288181305
step: 130, loss: 0.08515913784503937
step: 140, loss: 0.0625990480184555
step: 150, loss: 0.07523400336503983
step: 160, loss: 0.037282366305589676
step: 170, loss: 0.26965516805648804
step: 180, loss: 0.12731577455997467
step: 190, loss: 0.008973128162324429
step: 200, loss: 0.1229933574795723
step: 210, loss: 0.06395336240530014
step: 220, loss: 0.2507479190826416
step: 230, loss: 0.2261808216571808
step: 240, loss: 0.08073965460062027
step: 250, loss: 0.03367999196052551
step: 260, loss: 0.09872864931821823
step: 270, loss: 0.037544820457696915
step: 280, loss: 0.052364531904459
step: 290, loss: 0.03695635125041008
step: 300, loss: 0.0626295953989029
step: 310, loss: 0.16837626695632935
step: 320, loss: 0.1263977736234665
step: 330, loss: 0.04656391590833664
step: 340, loss: 0.08228420466184616
step: 350, loss: 0.0715625137090683
step: 360, loss: 0.022497661411762238
step: 370, loss: 0.07781954854726791
step: 380, loss: 0.051216308027505875
step: 390, loss: 0.1122487261891365
step: 400, loss: 0.31578704714775085
step: 410, loss: 0.0562257282435894
step: 420, loss: 0.059482648968696594
step: 430, loss: 0.15740707516670227
step: 440, loss: 0.04486233741044998
step: 450, loss: 0.01222827285528183
step: 460, loss: 0.014164031483232975

```

step: 470, loss: 0.1311779022216797
step: 480, loss: 0.05476957559585571
step: 490, loss: 0.031459566205739975
step: 500, loss: 0.07810099422931671
step: 510, loss: 0.06183811277151108
step: 520, loss: 0.1672077625989914
step: 530, loss: 0.009228629991412163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9203539823008848, f1=0.9130434782608695,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.081429123878479
step: 10, loss: 0.0770966187119484
step: 20, loss: 0.009426066651940346
step: 30, loss: 0.010710712522268295
step: 40, loss: 0.07976049184799194
step: 50, loss: 0.15055051445960999
step: 60, loss: 0.12604399025440216
step: 70, loss: 0.050763994455337524
step: 80, loss: 0.044467393308877945
step: 90, loss: 0.022672422230243683
step: 100, loss: 0.023452317342162132
step: 110, loss: 0.03800850361585617
step: 120, loss: 0.14865852892398834
step: 130, loss: 0.14924080669879913
step: 140, loss: 0.029922474175691605
step: 150, loss: 0.06508482247591019
step: 160, loss: 0.04253029450774193
step: 170, loss: 0.06178855150938034
step: 180, loss: 0.019600534811615944
step: 190, loss: 0.11905282735824585
step: 200, loss: 0.021203072741627693
step: 210, loss: 0.040428318083286285
step: 220, loss: 0.11932086944580078
step: 230, loss: 0.01234261691570282
step: 240, loss: 0.024128448218107224
step: 250, loss: 0.025565287098288536
step: 260, loss: 0.0048911115154623985
step: 270, loss: 0.1727438122034073
step: 280, loss: 0.009682704694569111

```

```

step: 290, loss: 0.014419362880289555
step: 300, loss: 0.06684468686580658
step: 310, loss: 0.021642234176397324
step: 320, loss: 0.1598019301891327
step: 330, loss: 0.03696636110544205
step: 340, loss: 0.03280016779899597
step: 350, loss: 0.0020712176337838173
step: 360, loss: 0.09613098204135895
step: 370, loss: 0.154923677444458
step: 380, loss: 0.05603000894188881
step: 390, loss: 0.04081781208515167
step: 400, loss: 0.1554276943206787
step: 410, loss: 0.01738899014890194
step: 420, loss: 0.032437440007925034
step: 430, loss: 0.019655155017971992
step: 440, loss: 0.16098110377788544
step: 450, loss: 0.01652863807976246
step: 460, loss: 0.055281903594732285
step: 470, loss: 0.11789113283157349
step: 480, loss: 0.4278623163700104
step: 490, loss: 0.03598913177847862
step: 500, loss: 0.21421024203300476
step: 510, loss: 0.02341023087501526
step: 520, loss: 0.08653060346841812
step: 530, loss: 0.15846602618694305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9233610341643583, f1=0.9145496535796767,
best_f1=0.9145496535796767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08192858099937439
step: 10, loss: 0.012852209620177746
step: 20, loss: 0.05879022553563118
step: 30, loss: 0.15031003952026367
step: 40, loss: 0.05729731172323227
step: 50, loss: 0.14374418556690216
step: 60, loss: 0.004919878672808409
step: 70, loss: 0.007329223677515984
step: 80, loss: 0.014228804036974907
step: 90, loss: 0.006476063746958971
step: 100, loss: 0.005565886851400137

```

```
step: 110, loss: 0.18270602822303772
step: 120, loss: 0.013093468733131886
step: 130, loss: 0.010943813249468803
step: 140, loss: 0.061406414955854416
step: 150, loss: 0.047629307955503464
step: 160, loss: 0.04836969077587128
step: 170, loss: 0.10984107106924057
step: 180, loss: 0.05336194485425949
step: 190, loss: 0.015134195797145367
step: 200, loss: 0.10683742165565491
step: 210, loss: 0.10260958969593048
step: 220, loss: 0.13006703555583954
step: 230, loss: 0.0842580795288086
step: 240, loss: 0.0020101284608244896
step: 250, loss: 0.14691632986068726
step: 260, loss: 0.010343853384256363
step: 270, loss: 0.011261414736509323
step: 280, loss: 0.09885483235120773
step: 290, loss: 0.013639459386467934
step: 300, loss: 0.026333525776863098
step: 310, loss: 0.031773779541254044
step: 320, loss: 0.005251922179013491
step: 330, loss: 0.0018414126243442297
step: 340, loss: 0.055172815918922424
step: 350, loss: 0.019319869577884674
step: 360, loss: 0.02489439956843853
step: 370, loss: 0.008472135290503502
step: 380, loss: 0.006699172779917717
step: 390, loss: 0.07104812562465668
step: 400, loss: 0.0507795512676239
step: 410, loss: 0.021121660247445107
step: 420, loss: 0.0965552031993866
step: 430, loss: 0.010561088100075722
step: 440, loss: 0.00310129439458251
step: 450, loss: 0.050726283341646194
step: 460, loss: 0.0844557136297226
step: 470, loss: 0.006339406128972769
step: 480, loss: 0.00309115182608366
step: 490, loss: 0.005404432769864798
step: 500, loss: 0.004630680661648512
step: 510, loss: 0.021675225347280502
step: 520, loss: 0.058281391859054565
step: 530, loss: 0.03674619644880295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9270307480495641, f1=0.9127640036730946,
best_f1=0.9127640036730946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013741725124418736
step: 10, loss: 0.012396445497870445
step: 20, loss: 0.0037293790373951197
step: 30, loss: 0.041524261236190796
step: 40, loss: 0.07631859928369522
step: 50, loss: 0.008462046273052692
step: 60, loss: 0.024099519476294518
step: 70, loss: 0.0023330780677497387
step: 80, loss: 0.005593698471784592
step: 90, loss: 0.08900821954011917
step: 100, loss: 0.0014598392881453037
step: 110, loss: 0.006773500237613916
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.0010543246753513813
step: 130, loss: 0.0008528122561983764
step: 140, loss: 0.005254766438156366
step: 150, loss: 0.006650274619460106
step: 160, loss: 0.12746275961399078
step: 170, loss: 0.0017574377125129104
step: 180, loss: 0.012548940256237984
step: 190, loss: 0.010294786654412746
step: 200, loss: 0.02823656052350998
step: 210, loss: 0.0345170795917511
step: 220, loss: 0.0023230223450809717
step: 230, loss: 0.015737980604171753
step: 240, loss: 0.013638894073665142
step: 250, loss: 0.12987929582595825
step: 260, loss: 0.0083996020257473
step: 270, loss: 0.0029385872185230255
step: 280, loss: 0.2489011138677597
step: 290, loss: 0.01087256707251072
step: 300, loss: 0.005807085428386927
step: 310, loss: 0.006374477408826351
step: 320, loss: 0.0209135003387928
step: 330, loss: 0.002568407217040658
step: 340, loss: 0.0028914224822074175
step: 350, loss: 0.005113308317959309
step: 360, loss: 0.08333540707826614
step: 370, loss: 0.02264793962240219
step: 380, loss: 0.002962330123409629
step: 390, loss: 0.008127373643219471

```

```

step: 400, loss: 0.0019510373240336776
step: 410, loss: 0.004560247529298067
step: 420, loss: 0.08693471550941467
step: 430, loss: 0.22392204403877258
step: 440, loss: 0.001808484666980803
step: 450, loss: 0.005149366799741983
step: 460, loss: 0.003206892404705286
step: 470, loss: 0.08095579594373703
step: 480, loss: 0.021083682775497437
step: 490, loss: 0.006354154087603092
step: 500, loss: 0.008793159388005733
step: 510, loss: 0.0029833002481609583
step: 520, loss: 0.1008371114730835
step: 530, loss: 0.022727899253368378
/usr/local/lib/python3.7/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=0.9263157894736843, f1=0.9186954524575103,
best_f1=0.9127640036730946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10283205658197403
step: 10, loss: 0.04788130149245262
step: 20, loss: 0.003030196065083146
step: 30, loss: 0.0009322733385488391
step: 40, loss: 0.03970618173480034
step: 50, loss: 0.05166054889559746
step: 60, loss: 0.05367046222090721
step: 70, loss: 0.1025557890534401
step: 80, loss: 0.0009813992073759437
step: 90, loss: 0.010463383980095387
step: 100, loss: 0.02208722196519375
step: 110, loss: 0.0009936020942404866
step: 120, loss: 0.0009783218847587705
step: 130, loss: 0.0005229557282291353
step: 140, loss: 0.006526585668325424
step: 150, loss: 0.0019262911519035697
step: 160, loss: 0.0011210368247702718
step: 170, loss: 0.028253396973013878
step: 180, loss: 0.0014067781157791615
step: 190, loss: 0.0006428808555938303
step: 200, loss: 0.004081569146364927
step: 210, loss: 0.01516580767929554
step: 220, loss: 0.0012654296588152647
step: 230, loss: 0.03663908317685127
step: 240, loss: 0.00941087119281292
step: 250, loss: 0.17828302085399628

```



```

step: 260, loss: 0.0027501985896378756
step: 270, loss: 0.0019013751298189163
step: 280, loss: 0.0018779539968818426
step: 290, loss: 0.1376510113477707
step: 300, loss: 0.016253652051091194
step: 310, loss: 0.015312116593122482
step: 320, loss: 0.008489600382745266
step: 330, loss: 0.0064019407145679
step: 340, loss: 0.004747437313199043
step: 350, loss: 0.00509376497939229
step: 360, loss: 0.02593894489109516
step: 370, loss: 0.02161380648612976
step: 380, loss: 0.0010484757367521524
step: 390, loss: 0.0002637388533912599
step: 400, loss: 0.005391017068177462
step: 410, loss: 0.0007221625419333577
step: 420, loss: 0.00047998438822105527
step: 430, loss: 0.0009814968798309565
step: 440, loss: 0.13164310157299042
step: 450, loss: 0.0024298946373164654
step: 460, loss: 0.0007178090163506567
step: 470, loss: 0.0039551761001348495
step: 480, loss: 0.0045216684229671955
step: 490, loss: 0.000824660062789917
step: 500, loss: 0.010637352243065834
step: 510, loss: 0.1328793317079544
step: 520, loss: 0.06354048103094101
step: 530, loss: 0.00036740428186021745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9314179796107508, f1=0.9167054443927408,
best_f1=0.9167054443927408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003090861951932311
step: 10, loss: 0.0001137877261498943
step: 20, loss: 0.11200731992721558
step: 30, loss: 0.0005796621553599834
step: 40, loss: 0.0004432992427609861
step: 50, loss: 0.049609243869781494
step: 60, loss: 0.00043250469025224447
step: 70, loss: 0.005684811156243086

```

step: 80, loss: 0.0028492542915046215
step: 90, loss: 0.00043692972394637764
step: 100, loss: 0.004202053416520357
step: 110, loss: 0.0007122750976122916
step: 120, loss: 0.0001243191072717309
step: 130, loss: 0.0003438019484747201
step: 140, loss: 0.004051670897752047
step: 150, loss: 0.00282890722155571
step: 160, loss: 0.0018457922851666808
step: 170, loss: 0.0006167714600451291
step: 180, loss: 0.000853032513987273
step: 190, loss: 0.000968008884228766
step: 200, loss: 0.0001806875370675698
step: 210, loss: 0.004453637637197971
step: 220, loss: 0.005129234399646521
step: 230, loss: 0.0016872209962457418
step: 240, loss: 0.11644361168146133
step: 250, loss: 0.0009873538510873914
step: 260, loss: 0.0002976394898723811
step: 270, loss: 0.06085243821144104
step: 280, loss: 0.008302122354507446
step: 290, loss: 0.0014793741283938289
step: 300, loss: 0.009552095085382462
step: 310, loss: 0.007787294220179319
step: 320, loss: 0.002331023570150137
step: 330, loss: 0.0005841681850142777
step: 340, loss: 0.04079679027199745
step: 350, loss: 0.013385222293436527
step: 360, loss: 0.009089267812669277
step: 370, loss: 0.12395334988832474
step: 380, loss: 0.0002501348499208689
step: 390, loss: 0.0037438273429870605
step: 400, loss: 0.0005714892176911235
step: 410, loss: 0.0004295755352359265
step: 420, loss: 0.0011009472655132413
step: 430, loss: 0.0004626075678970665
step: 440, loss: 0.0020562950521707535
step: 450, loss: 0.00020726806542370468
step: 460, loss: 0.0002640667953528464
step: 470, loss: 0.0007910013664513826
step: 480, loss: 0.003915733192116022
step: 490, loss: 0.011225978843867779
step: 500, loss: 0.0005078435060568154
step: 510, loss: 0.006529238540679216
step: 520, loss: 0.004329431802034378
step: 530, loss: 0.005098668858408928
/usr/local/lib/python3.7/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=0.9281767955801105, f1=0.9150507848568791,
best_f1=0.9167054443927408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011133011430501938
step: 10, loss: 0.0018755890196189284
step: 20, loss: 0.007855688221752644
step: 30, loss: 0.001422672881744802
step: 40, loss: 0.0037954135332256556
step: 50, loss: 0.0006821380229666829
step: 60, loss: 0.0028417001012712717
step: 70, loss: 0.001185975386761129
step: 80, loss: 0.01141010969877243
step: 90, loss: 0.0001397771993651986
step: 100, loss: 0.00013431163097266108
step: 110, loss: 0.003679753514006734
step: 120, loss: 0.0005938736139796674
step: 130, loss: 0.04421176388859749
step: 140, loss: 0.00010762065357994288
step: 150, loss: 0.0004542575334198773
step: 160, loss: 0.0009130289545282722
step: 170, loss: 0.0008612402598373592
step: 180, loss: 0.0006523763295263052
step: 190, loss: 0.09192227572202682
step: 200, loss: 0.055169254541397095
step: 210, loss: 0.0007490451098419726
step: 220, loss: 0.0004515148757491261
step: 230, loss: 0.047350965440273285
step: 240, loss: 0.0002069797774311155
step: 250, loss: 0.00033012873609550297
step: 260, loss: 0.00012150712427683175
step: 270, loss: 0.00038629258051514626
step: 280, loss: 6.515807763207704e-05
step: 290, loss: 0.00014889618614688516
step: 300, loss: 0.0015902987215667963
step: 310, loss: 0.0015864840243011713
step: 320, loss: 0.0003300172393210232
step: 330, loss: 0.0012878805864602327
step: 340, loss: 0.03696996346116066
step: 350, loss: 0.03070257417857647
step: 360, loss: 0.0023271245881915092
step: 370, loss: 0.002094893017783761
step: 380, loss: 0.0003117906744591892
step: 390, loss: 0.00014468458539340645
step: 400, loss: 0.00251327664591372
step: 410, loss: 0.001546178595162928
```

```
step: 420, loss: 0.0005504846922121942
step: 430, loss: 0.00027998819132335484
step: 440, loss: 0.0006864204769954085
step: 450, loss: 0.003088476601988077
step: 460, loss: 0.09335466474294662
step: 470, loss: 0.00027623947244137526
step: 480, loss: 0.13489557802677155
step: 490, loss: 0.030275817960500717
step: 500, loss: 0.0018497315468266606
step: 510, loss: 0.0014008505968376994
step: 520, loss: 0.0015548813389614224
step: 530, loss: 0.003536381060257554
/usr/local/lib/python3.7/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=0.9285714285714286, f1=0.9133709981167608,
best_f1=0.9167054443927408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005302937352098525
step: 10, loss: 0.038184743374586105
step: 20, loss: 0.0004657669342122972
step: 30, loss: 0.0029263091273605824
step: 40, loss: 0.0004222997813485563
step: 50, loss: 0.00568418949842453
step: 60, loss: 0.0016480403719469905
step: 70, loss: 0.0010740096913650632
step: 80, loss: 0.0034194595646113157
step: 90, loss: 0.0007269702618941665
step: 100, loss: 0.0031088439282029867
step: 110, loss: 0.00018161247135140002
step: 120, loss: 8.437027281615883e-05
step: 130, loss: 0.00011394873581593856
step: 140, loss: 0.01866840198636055
step: 150, loss: 0.014982235617935658
step: 160, loss: 0.0013409459497779608
step: 170, loss: 0.0246792770922184
step: 180, loss: 0.0006646759575232863
step: 190, loss: 0.0028249239549040794
step: 200, loss: 0.0023988159373402596
step: 210, loss: 8.2083344750572e-05
step: 220, loss: 0.0019343794556334615
step: 230, loss: 5.43209862371441e-05
step: 240, loss: 0.003986597526818514
step: 250, loss: 5.1013583288295195e-05
step: 260, loss: 6.350220792228356e-05
step: 270, loss: 0.0010913914302363992
```

```

step: 280, loss: 0.012356990948319435
step: 290, loss: 9.884306928142905e-05
step: 300, loss: 0.001536583062261343
step: 310, loss: 0.0025459984317421913
step: 320, loss: 0.0006389853660948575
step: 330, loss: 0.00012783441343344748
step: 340, loss: 0.005213044583797455
step: 350, loss: 0.00010025369556387886
step: 360, loss: 0.00011507069575600326
step: 370, loss: 0.011126790195703506
step: 380, loss: 0.0013818728039041162
step: 390, loss: 0.1355276107788086
step: 400, loss: 0.0011694746790453792
step: 410, loss: 4.754019755637273e-05
step: 420, loss: 0.004013998433947563
step: 430, loss: 0.011272349394857883
step: 440, loss: 0.0010728357592597604
step: 450, loss: 0.00010566256241872907
step: 460, loss: 7.424199429806322e-05
step: 470, loss: 0.00012820739357266575
step: 480, loss: 4.973830436938442e-05
step: 490, loss: 0.0746549591422081
step: 500, loss: 0.0013474924489855766
step: 510, loss: 0.0006338901002891362
step: 520, loss: 0.0020591169595718384
step: 530, loss: 0.00010105586261488497
/usr/local/lib/python3.7/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=0.9288702928870293, f1=0.9245719574271172,
best_f1=0.9167054443927408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011356055038049817
step: 10, loss: 5.6746284826658666e-05
step: 20, loss: 0.005594408139586449
step: 30, loss: 9.540725295664743e-05
step: 40, loss: 5.55664264538791e-05
step: 50, loss: 0.0021827761083841324
step: 60, loss: 0.0002145320177078247
step: 70, loss: 0.00019905739463865757
step: 80, loss: 0.0017541530542075634
step: 90, loss: 4.204975994070992e-05
step: 100, loss: 9.164397488348186e-05
step: 110, loss: 5.125607276568189e-05
step: 120, loss: 0.000164354030857794
step: 130, loss: 0.00013604322157334536

```

```

step: 140, loss: 0.0005845591076649725
step: 150, loss: 4.811240069102496e-05
step: 160, loss: 0.00030831515323370695
step: 170, loss: 0.019672060385346413
step: 180, loss: 5.290669287205674e-05
step: 190, loss: 7.536332122981548e-05
step: 200, loss: 4.939422433380969e-05
step: 210, loss: 0.012853272259235382
step: 220, loss: 0.006997366901487112
step: 230, loss: 3.112769627477974e-05
step: 240, loss: 0.00026196156977675855
step: 250, loss: 0.0003480963932815939
step: 260, loss: 0.001621865900233388
step: 270, loss: 2.745485289779026e-05
step: 280, loss: 3.431998266023584e-05
step: 290, loss: 0.04732677340507507
step: 300, loss: 5.056178997620009e-05
step: 310, loss: 0.03131946921348572
step: 320, loss: 0.0004889533156529069
step: 330, loss: 0.0003853460948448628
step: 340, loss: 0.01941317319869995
step: 350, loss: 0.007106594741344452
step: 360, loss: 0.0001884922239696607
step: 370, loss: 0.0032950108870863914
step: 380, loss: 0.0002589064242783934
step: 390, loss: 7.795955752953887e-05
step: 400, loss: 0.0018600283656269312
step: 410, loss: 0.00022858740703668445
step: 420, loss: 6.077413127059117e-05
step: 430, loss: 0.0004036214086227119
step: 440, loss: 0.002091895090416074
step: 450, loss: 5.128681368660182e-05
step: 460, loss: 0.00017894922348205
step: 470, loss: 3.40205297106877e-05
step: 480, loss: 4.737864583148621e-05
step: 490, loss: 0.0005619236035272479
step: 500, loss: 0.0033593778498470783
step: 510, loss: 0.13445761799812317
step: 520, loss: 3.928871592506766e-05
step: 530, loss: 0.00011031256144633517
/usr/local/lib/python3.7/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=0.9291412482402628, f1=0.9145823501651722,
best_f1=0.9167054443927408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.00015641718346159905
step: 10, loss: 3.880063013639301e-05
step: 20, loss: 2.52531190199079e-05
step: 30, loss: 0.00010418256715638563
step: 40, loss: 0.0001414556900272146
step: 50, loss: 0.015865743160247803
step: 60, loss: 0.0002524346054997295
step: 70, loss: 0.00012078888539690524
step: 80, loss: 3.37601377395913e-05
step: 90, loss: 4.3518160964595154e-05
step: 100, loss: 0.00011815902689704672
step: 110, loss: 0.00212428392842412
step: 120, loss: 0.001906525925733149
step: 130, loss: 0.0008432078175246716
step: 140, loss: 0.0028347105253487825
step: 150, loss: 0.02356254868209362
step: 160, loss: 4.0840968722477555e-05
step: 170, loss: 7.837721204850823e-05
step: 180, loss: 0.00012789072934538126
step: 190, loss: 6.703607505187392e-05
step: 200, loss: 0.022002598270773888
step: 210, loss: 7.044304948067293e-05
step: 220, loss: 0.0017488619778305292
step: 230, loss: 0.0005895736394450068
step: 240, loss: 0.001243136590346694
step: 250, loss: 7.577969518024474e-05
step: 260, loss: 0.006430424749851227
step: 270, loss: 0.0015182122588157654
step: 280, loss: 0.0018289723666384816
step: 290, loss: 9.975412103813142e-05
step: 300, loss: 3.8021364161977544e-05
step: 310, loss: 5.028045052313246e-05
step: 320, loss: 0.014502095058560371
step: 330, loss: 0.0003471876843832433
step: 340, loss: 0.00014540087431669235
step: 350, loss: 0.00013052576105110347
step: 360, loss: 0.001152741489931941
step: 370, loss: 0.001988017465919256
step: 380, loss: 0.0005836626514792442
step: 390, loss: 0.0006336785736493766
step: 400, loss: 0.0013496748870238662
step: 410, loss: 0.0011954924557358027
step: 420, loss: 9.570588736096397e-05
step: 430, loss: 2.069370384560898e-05
step: 440, loss: 0.0003624742676038295
step: 450, loss: 9.172748832497746e-05
step: 460, loss: 4.718609125120565e-05
step: 470, loss: 0.0001096693886211142

```

step: 480, loss: 0.00022698118118569255
step: 490, loss: 0.000154231209307909
step: 500, loss: 0.0025146761909127235
step: 510, loss: 4.121322854189202e-05
step: 520, loss: 0.02947952412068844
step: 530, loss: 0.0012836025562137365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9334574220567706, f1=0.9185667752442996,
best_f1=0.9185667752442996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040292413905262947
step: 10, loss: 0.00023234930995386094
step: 20, loss: 0.000152146618347615
step: 30, loss: 0.0007286449545063078
step: 40, loss: 0.000305371533613652
step: 50, loss: 7.149034354370087e-05
step: 60, loss: 0.009597246535122395
step: 70, loss: 0.0003835636889562011
step: 80, loss: 2.6526933652348816e-05
step: 90, loss: 4.602083936333656e-05
step: 100, loss: 0.0005015283823013306
step: 110, loss: 4.9802911235019565e-05
step: 120, loss: 0.006367234513163567
step: 130, loss: 5.8086890931008384e-05
step: 140, loss: 0.00014301702321972698
step: 150, loss: 2.5908431780408137e-05
step: 160, loss: 4.7384492063429207e-05
step: 170, loss: 0.00016734884411562234
step: 180, loss: 0.00017432124877814204
step: 190, loss: 0.0006187880644574761
step: 200, loss: 0.002979251090437174
step: 210, loss: 0.0004970841691829264
step: 220, loss: 0.0002874744532164186
step: 230, loss: 0.0003511066024657339
step: 240, loss: 0.0001830561232054606
step: 250, loss: 2.045893961621914e-05
step: 260, loss: 7.249849295476452e-05
step: 270, loss: 0.000596950005274266
step: 280, loss: 0.00018585534417070448
step: 290, loss: 6.931180541869253e-05

```



```

step: 300, loss: 3.78943259420339e-05
step: 310, loss: 3.0024859370314516e-05
step: 320, loss: 8.10603960417211e-05
step: 330, loss: 0.0002489872567821294
step: 340, loss: 3.568635293049738e-05
step: 350, loss: 5.017878356738947e-05
step: 360, loss: 1.975091618078295e-05
step: 370, loss: 0.00026873135357163846
step: 380, loss: 0.0025772824883461
step: 390, loss: 6.228303391253576e-05
step: 400, loss: 0.0005104951560497284
step: 410, loss: 0.000579186889808625
step: 420, loss: 9.313254849985242e-05
step: 430, loss: 4.268114207661711e-05
step: 440, loss: 6.754988135071471e-05
step: 450, loss: 5.313456495059654e-05
step: 460, loss: 0.003044726559892297
step: 470, loss: 5.414076440501958e-05
step: 480, loss: 4.831537444260903e-05
step: 490, loss: 4.432143396115862e-05
step: 500, loss: 0.0002863148401957005
step: 510, loss: 4.963629908161238e-05
step: 520, loss: 2.608762588351965e-05
step: 530, loss: 2.016828875639476e-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=0.9263256687001408, f1=0.9128591615638247,
best_f1=0.9185667752442996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.269777178298682e-05
step: 10, loss: 1.851438355515711e-05
step: 20, loss: 2.404582301096525e-05
step: 30, loss: 0.00012398953549563885
step: 40, loss: 3.0430393962888047e-05
step: 50, loss: 0.00028128683334216475
step: 60, loss: 0.0004539405053947121
step: 70, loss: 5.743668953073211e-05
step: 80, loss: 4.216551678837277e-05
step: 90, loss: 5.166686241864227e-05
step: 100, loss: 0.00012983032502233982
step: 110, loss: 4.753219764097594e-05
step: 120, loss: 1.9240749679738656e-05
step: 130, loss: 3.5039378417422995e-05
step: 140, loss: 0.0002502024872228503
step: 150, loss: 9.766741277417168e-05

```

```

step: 160, loss: 2.3699607481830753e-05
step: 170, loss: 2.4157950974768028e-05
step: 180, loss: 2.870177740987856e-05
step: 190, loss: 3.6754761822521687e-05
step: 200, loss: 3.4743890864774585e-05
step: 210, loss: 3.493634721962735e-05
step: 220, loss: 3.9067686884664e-05
step: 230, loss: 1.6711406715330668e-05
step: 240, loss: 0.01064245868474245
step: 250, loss: 1.803752275009174e-05
step: 260, loss: 2.0891140593448654e-05
step: 270, loss: 3.021497468580492e-05
step: 280, loss: 0.00035355513682588935
step: 290, loss: 2.6638697818270884e-05
step: 300, loss: 2.2720048946212046e-05
step: 310, loss: 0.0002141878940165043
step: 320, loss: 2.6518826416577213e-05
step: 330, loss: 5.3929608839098364e-05
step: 340, loss: 1.8097171050612815e-05
step: 350, loss: 0.00010139232472283766
step: 360, loss: 0.00037586508551612496
step: 370, loss: 6.187500548548996e-05
step: 380, loss: 3.325056241010316e-05
step: 390, loss: 3.288225343567319e-05
step: 400, loss: 3.409056080272421e-05
step: 410, loss: 2.219105044787284e-05
step: 420, loss: 0.015401928685605526
step: 430, loss: 9.28933295654133e-05
step: 440, loss: 2.6325538783567026e-05
step: 450, loss: 6.162809586385265e-05
step: 460, loss: 2.0786743334610946e-05
step: 470, loss: 2.1110501620569266e-05
step: 480, loss: 0.0002688126696739346
step: 490, loss: 4.7304059989983216e-05
step: 500, loss: 6.20575956418179e-05
step: 510, loss: 0.004495192784816027
step: 520, loss: 2.9483884645742364e-05
step: 530, loss: 5.492181298905052e-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=0.9321016166281756, f1=0.9155844155844156,
best_f1=0.9185667752442996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008841765113174915
step: 10, loss: 2.2813188479631208e-05

```

step: 20, loss: 1.7281372493016534e-05
step: 30, loss: 5.014983253204264e-05
step: 40, loss: 2.0354549633339047e-05
step: 50, loss: 0.00014963233843445778
step: 60, loss: 6.794249202357605e-05
step: 70, loss: 1.471473206038354e-05
step: 80, loss: 1.861126656876877e-05
step: 90, loss: 0.0017979200929403305
step: 100, loss: 0.0007283795857802033
step: 110, loss: 0.00016250237240456045
step: 120, loss: 1.6979638530756347e-05
step: 130, loss: 1.873768815130461e-05
step: 140, loss: 3.878030111081898e-05
step: 150, loss: 4.628744864021428e-05
step: 160, loss: 2.0521969418041408e-05
step: 170, loss: 0.0015143360942602158
step: 180, loss: 3.0072864319663495e-05
step: 190, loss: 0.00020420343207661062
step: 200, loss: 3.7711641198256984e-05
step: 210, loss: 2.9200986318755895e-05
step: 220, loss: 1.8521786842029542e-05
step: 230, loss: 1.94826479855692e-05
step: 240, loss: 3.0209443139028735e-05
step: 250, loss: 1.986272764042951e-05
step: 260, loss: 1.9929644622607157e-05
step: 270, loss: 4.0334693039767444e-05
step: 280, loss: 3.512686089379713e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 290, loss: 0.0014811345608904958
step: 300, loss: 2.198621405113954e-05
step: 310, loss: 0.00028076075250282884
step: 320, loss: 0.000750821374822408
step: 330, loss: 3.985271177953109e-05
step: 340, loss: 0.00010774791007861495
step: 350, loss: 2.17699671338778e-05
step: 360, loss: 2.0671192032750696e-05
step: 370, loss: 0.00011530109622981399
step: 380, loss: 7.971916056703776e-05
step: 390, loss: 0.0025030660908669233
step: 400, loss: 4.6261746319942176e-05
step: 410, loss: 2.150939508283045e-05
step: 420, loss: 2.651123759278562e-05
step: 430, loss: 0.0013806408969685435
step: 440, loss: 0.0001427509996574372
step: 450, loss: 0.007152384612709284
step: 460, loss: 5.9093912568641827e-05
step: 470, loss: 1.2859572962042876e-05
step: 480, loss: 1.534054354124237e-05

```

step: 490, loss: 2.220194073743187e-05
step: 500, loss: 1.183143558591837e-05
step: 510, loss: 2.060781480395235e-05
step: 520, loss: 5.632108877762221e-05
step: 530, loss: 1.2628620424948167e-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=0.9324639031206335, f1=0.9196261682242991,
best_f1=0.9185667752442996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001093925442546606
step: 10, loss: 1.3194833627494518e-05
step: 20, loss: 2.4593058697064407e-05
step: 30, loss: 2.880897773138713e-05
step: 40, loss: 3.9078600821085274e-05
step: 50, loss: 0.004150783643126488
step: 60, loss: 2.308431976416614e-05
step: 70, loss: 3.587009268812835e-05
step: 80, loss: 3.8996866351226345e-05
step: 90, loss: 1.7355885574943386e-05
step: 100, loss: 1.560877717565745e-05
step: 110, loss: 3.382941577001475e-05
step: 120, loss: 4.763670949614607e-05
step: 130, loss: 0.002512169536203146
step: 140, loss: 2.3061462343321182e-05
step: 150, loss: 1.8394870494375937e-05
step: 160, loss: 1.7694839698378928e-05
step: 170, loss: 1.84696455107769e-05
step: 180, loss: 2.295378544658888e-05
step: 190, loss: 0.0001655852101976052
step: 200, loss: 5.4976560932118446e-05
step: 210, loss: 3.126760930172168e-05
step: 220, loss: 1.7571834177942947e-05
step: 230, loss: 2.5047722374438308e-05
step: 240, loss: 0.00019916649034712464
step: 250, loss: 1.408515981893288e-05
step: 260, loss: 1.6703956134733744e-05
step: 270, loss: 1.3500295608537272e-05
step: 280, loss: 1.2595091902767308e-05
step: 290, loss: 0.006788522470742464
step: 300, loss: 5.8075354900211096e-05
step: 310, loss: 2.2712600184604526e-05
step: 320, loss: 0.0003857404226437211
step: 330, loss: 7.639022078365088e-05
step: 340, loss: 1.746759699017275e-05

```

```

step: 350, loss: 0.0006706771091558039
step: 360, loss: 2.9364178772084415e-05
step: 370, loss: 2.5196266506100073e-05
step: 380, loss: 2.5661884137662128e-05
step: 390, loss: 0.004570521414279938
step: 400, loss: 0.0008401665836572647
step: 410, loss: 1.7575495803612284e-05
step: 420, loss: 1.7087575542973354e-05
step: 430, loss: 4.059209459228441e-05
step: 440, loss: 4.9614198360359296e-05
step: 450, loss: 5.95318888372276e-05
step: 460, loss: 3.92750880564563e-05
step: 470, loss: 1.8689381249714643e-05
step: 480, loss: 3.8747584767406806e-05
step: 490, loss: 2.26339452638058e-05
step: 500, loss: 1.776168392098043e-05
step: 510, loss: 3.7314453948056325e-05
step: 520, loss: 2.9519665986299515e-05
step: 530, loss: 2.670764115464408e-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=0.9307116104868914, f1=0.9172932330827067,
best_f1=0.9185667752442996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.0641407900257036e-05
step: 10, loss: 2.584846515674144e-05
step: 20, loss: 2.532656981202308e-05
step: 30, loss: 1.5489595170947723e-05
step: 40, loss: 0.038014817982912064
step: 50, loss: 0.00040192570304498076
step: 60, loss: 1.2617436368600465e-05
step: 70, loss: 0.00039999111322686076
step: 80, loss: 0.00016431871335953474
step: 90, loss: 1.9181170500814915e-05
step: 100, loss: 4.351049938122742e-05
step: 110, loss: 0.0012715024640783668
step: 120, loss: 0.00011471117613837123
step: 130, loss: 0.029110126197338104
step: 140, loss: 0.00026043131947517395
step: 150, loss: 2.7681346182362176e-05
step: 160, loss: 2.1240830392343923e-05
step: 170, loss: 1.4223009202396497e-05
step: 180, loss: 1.8781822291202843e-05
step: 190, loss: 2.0738081730087288e-05
step: 200, loss: 0.06945338845252991

```

```

step: 210, loss: 0.00045652661356143653
step: 220, loss: 0.00022820101003162563
step: 230, loss: 1.9266828530817293e-05
step: 240, loss: 1.4379463209479582e-05
step: 250, loss: 1.9419388991082087e-05
step: 260, loss: 1.6227082596742548e-05
step: 270, loss: 1.740803236316424e-05
step: 280, loss: 2.264176328026224e-05
step: 290, loss: 1.2263541066204198e-05
step: 300, loss: 2.3673230316489935e-05
step: 310, loss: 2.549108103266917e-05
step: 320, loss: 1.4401713997358456e-05
step: 330, loss: 1.6349982615793124e-05
step: 340, loss: 1.3839311577612534e-05
step: 350, loss: 1.1287541383353528e-05
step: 360, loss: 0.0002523482544347644
step: 370, loss: 1.1544589142431505e-05
step: 380, loss: 2.0037883587065153e-05
step: 390, loss: 2.4037657567532733e-05
step: 400, loss: 1.5687021004850976e-05
step: 410, loss: 1.444646386516979e-05
step: 420, loss: 6.515517452498898e-05
step: 430, loss: 6.199078779900447e-05
step: 440, loss: 3.275583731010556e-05
step: 450, loss: 1.2177850294392556e-05
step: 460, loss: 1.4722151718160603e-05
step: 470, loss: 0.0074762399308383465
step: 480, loss: 1.3895161828259006e-05
step: 490, loss: 1.6022238924051635e-05
step: 500, loss: 2.215362292190548e-05
step: 510, loss: 2.004865564231295e-05
step: 520, loss: 1.8097058273269795e-05
step: 530, loss: 3.0324761610245332e-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=0.9337042188224385, f1=0.9215777262180975,
best_f1=0.9215777262180975
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 [00:24, 229.96it/s]
load_f1 = 0.929889298892989
real_f1 = 0.9296551724137931
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 240.87it/s]

```

3.1.4 Amazon-Google - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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


```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
```

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

```
step: 0, loss: 0.5551915764808655
step: 10, loss: 0.36416417360305786
step: 20, loss: 0.38340309262275696
step: 30, loss: 0.3306896388530731
step: 40, loss: 0.19192223250865936
step: 50, loss: 0.4726487696170807
step: 60, loss: 0.2644343376159668
step: 70, loss: 0.22812418639659882
step: 80, loss: 0.20013248920440674
step: 90, loss: 0.4159070551395416
step: 100, loss: 0.4963264763355255
step: 110, loss: 0.2339165359735489
step: 120, loss: 0.20973162353038788
step: 130, loss: 0.3025988042354584
step: 140, loss: 0.25363776087760925
step: 150, loss: 0.2713795006275177
step: 160, loss: 0.40856868028640747
step: 170, loss: 0.32619622349739075
step: 180, loss: 0.1723475158214569
step: 190, loss: 0.19059328734874725
step: 200, loss: 0.36307328939437866
step: 210, loss: 0.1754404902458191
```

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

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

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

```
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=0.5199999999999999, f1=0.5083333333333333,
best_f1=0.5083333333333333
```

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

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

```
step: 0, loss: 0.12205299735069275
step: 10, loss: 0.25114893913269043
step: 20, loss: 0.2540398836135864
step: 30, loss: 0.1610337793827057
step: 40, loss: 0.2453099489212036
```

```

step: 50, loss: 0.20988619327545166
step: 60, loss: 0.43217381834983826
step: 70, loss: 0.12975138425827026
step: 80, loss: 0.21464264392852783
step: 90, loss: 0.08534016460180283
step: 100, loss: 0.02505008690059185
step: 110, loss: 0.11289355903863907
step: 120, loss: 0.10824297368526459
step: 130, loss: 0.051018983125686646
step: 140, loss: 0.25588488578796387
step: 150, loss: 0.27184754610061646
step: 160, loss: 0.20933949947357178
step: 170, loss: 0.1804114580154419
step: 180, loss: 0.14166909456253052
step: 190, loss: 0.20499394834041595
step: 200, loss: 0.05524178594350815
step: 210, loss: 0.13527017831802368
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5658153241650296, f1=0.5919661733615221,
best_f1=0.5919661733615221
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05138922482728958
step: 10, loss: 0.17755332589149475
step: 20, loss: 0.08654891699552536
step: 30, loss: 0.26886147260665894
step: 40, loss: 0.09644395858049393
step: 50, loss: 0.17619308829307556
step: 60, loss: 0.1877298802137375
step: 70, loss: 0.15913666784763336
step: 80, loss: 0.17701129615306854
step: 90, loss: 0.0888882577419281
step: 100, loss: 0.1713719367980957
step: 110, loss: 0.15155619382858276
step: 120, loss: 0.08051116019487381
step: 130, loss: 0.15552367269992828
step: 140, loss: 0.06754716485738754
step: 150, loss: 0.12210585176944733
step: 160, loss: 0.030220001935958862
step: 170, loss: 0.10122649371623993
step: 180, loss: 0.042958907783031464

```

```

step: 190, loss: 0.19844616949558258
step: 200, loss: 0.08326489478349686
step: 210, loss: 0.14194892346858978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.5714285714285714, f1=0.5934959349593495,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1542309671640396
step: 10, loss: 0.0854736790060997
step: 20, loss: 0.019960595294833183
step: 30, loss: 0.07612097263336182
step: 40, loss: 0.045497994869947433
step: 50, loss: 0.20129218697547913
step: 60, loss: 0.10906577855348587
step: 70, loss: 0.245890811085701
step: 80, loss: 0.08438445627689362
step: 90, loss: 0.040821854025125504
step: 100, loss: 0.1813330352306366
step: 110, loss: 0.12989285588264465
step: 120, loss: 0.0923689752817154
step: 130, loss: 0.07475404441356659
step: 140, loss: 0.2172023504972458
step: 150, loss: 0.0657808855175972
step: 160, loss: 0.04639579355716705
step: 170, loss: 0.1360224187374115
step: 180, loss: 0.19850406050682068
step: 190, loss: 0.049174509942531586
step: 200, loss: 0.09391710162162781
step: 210, loss: 0.2609499990940094
/usr/local/lib/python3.7/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=0.5599999999999999, f1=0.5333333333333333,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1238248273730278
step: 10, loss: 0.018398895859718323
step: 20, loss: 0.09602808952331543

```

```

step: 30, loss: 0.026041196659207344
step: 40, loss: 0.008704275824129581
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.0659855306148529
step: 60, loss: 0.018973637372255325
step: 70, loss: 0.0543525367975235
step: 80, loss: 0.017914127558469772
step: 90, loss: 0.11523822695016861
step: 100, loss: 0.06264089047908783
step: 110, loss: 0.027585182338953018
step: 120, loss: 0.20362356305122375
step: 130, loss: 0.040515121072530746
step: 140, loss: 0.10668452829122543
step: 150, loss: 0.06454780697822571
step: 160, loss: 0.08257825672626495
step: 170, loss: 0.0653235912322998
step: 180, loss: 0.0547674186527729
step: 190, loss: 0.065955251455307
step: 200, loss: 0.01863657310605049
step: 210, loss: 0.020304203033447266
/usr/local/lib/python3.7/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=0.5649717514124294, f1=0.5338809034907598,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.048707108944654465
step: 10, loss: 0.027209747582674026
step: 20, loss: 0.012136539444327354
step: 30, loss: 0.024311158806085587
step: 40, loss: 0.004830789286643267
step: 50, loss: 0.016491398215293884
step: 60, loss: 0.017241282388567924
step: 70, loss: 0.015818089246749878
step: 80, loss: 0.04904545843601227
step: 90, loss: 0.027014251798391342
step: 100, loss: 0.0037765265442430973
step: 110, loss: 0.016036352142691612
step: 120, loss: 0.05033458396792412
step: 130, loss: 0.043546296656131744
step: 140, loss: 0.18421897292137146
step: 150, loss: 0.008497768081724644
step: 160, loss: 0.060127776116132736
step: 170, loss: 0.0769752711057663
step: 180, loss: 0.0043366351164877415
step: 190, loss: 0.2375979870557785

```

```

step: 200, loss: 0.01999126374721527
step: 210, loss: 0.010217485949397087
/usr/local/lib/python3.7/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=0.5397412199630314, f1=0.5333333333333333,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034827864728868008
step: 10, loss: 0.005292970687150955
step: 20, loss: 0.005281167104840279
step: 30, loss: 0.04004019871354103
step: 40, loss: 0.01088854018598795
step: 50, loss: 0.058956120163202286
step: 60, loss: 0.04303310438990593
step: 70, loss: 0.00806682463735342
step: 80, loss: 0.07261871546506882
step: 90, loss: 0.008427497930824757
step: 100, loss: 0.0033889522310346365
step: 110, loss: 0.01994786038994789
step: 120, loss: 0.04819241538643837
step: 130, loss: 0.08495964109897614
step: 140, loss: 0.007141214329749346
step: 150, loss: 0.023359326645731926
step: 160, loss: 0.0028899647295475006
step: 170, loss: 0.009519575163722038
step: 180, loss: 0.004079267382621765
step: 190, loss: 0.19401177763938904
step: 200, loss: 0.007923382334411144
step: 210, loss: 0.006822921335697174
/usr/local/lib/python3.7/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=0.5079365079365079, f1=0.4717444717444717,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026422932278364897
step: 10, loss: 0.019980384036898613
step: 20, loss: 0.010392076335847378
step: 30, loss: 0.0028988714329898357
step: 40, loss: 0.011924122460186481
step: 50, loss: 0.0039145792834460735
step: 60, loss: 0.03603003919124603
step: 70, loss: 0.00787327066063881

```

```

step: 80, loss: 0.031230585649609566
step: 90, loss: 0.013288953341543674
step: 100, loss: 0.11700829863548279
step: 110, loss: 0.017232730984687805
step: 120, loss: 0.0005588305066339672
step: 130, loss: 0.001334058353677392
step: 140, loss: 0.0796334370970726
step: 150, loss: 0.0069840336218476295
step: 160, loss: 0.034291740506887436
step: 170, loss: 0.10646801441907883
step: 180, loss: 0.1267383098602295
step: 190, loss: 0.050275348126888275
step: 200, loss: 0.0362812764942646
step: 210, loss: 0.03178155794739723
/usr/local/lib/python3.7/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=0.5502183406113538, f1=0.5081585081585082,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01120944507420063
step: 10, loss: 0.010283111594617367
step: 20, loss: 0.008419663645327091
step: 30, loss: 0.0033202795311808586
step: 40, loss: 0.1265724003314972
step: 50, loss: 0.0023433996830135584
step: 60, loss: 0.008583572693169117
step: 70, loss: 0.0007090627332217991
step: 80, loss: 0.0021378423552960157
step: 90, loss: 0.00031856700661592185
step: 100, loss: 0.000515105901286006
step: 110, loss: 0.008725438266992569
step: 120, loss: 0.00221979315392673
step: 130, loss: 0.04093186557292938
step: 140, loss: 0.010200290009379387
step: 150, loss: 0.05152691900730133
step: 160, loss: 0.0024333642795681953
step: 170, loss: 0.013159281574189663
step: 180, loss: 0.006627574563026428
step: 190, loss: 0.0005321530043147504
step: 200, loss: 0.010878045111894608
step: 210, loss: 0.002440195996314287
/usr/local/lib/python3.7/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=0.530701754385965, f1=0.5136363636363637,

```

```

best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02668123133480549
step: 10, loss: 0.005851753056049347
step: 20, loss: 0.0067029413767158985
step: 30, loss: 0.10214225947856903
step: 40, loss: 0.0009106603101827204
step: 50, loss: 0.01795695163309574
step: 60, loss: 0.006102082319557667
step: 70, loss: 0.008945068344473839
step: 80, loss: 0.0048091961070895195
step: 90, loss: 0.01258071232587099
step: 100, loss: 0.005571307148784399
step: 110, loss: 0.0002724274236243218
step: 120, loss: 0.0005706861265935004
step: 130, loss: 0.03368373215198517
step: 140, loss: 0.002808435121551156
step: 150, loss: 0.0019910959526896477
step: 160, loss: 0.004475272260606289
step: 170, loss: 0.0029419991187751293
step: 180, loss: 0.029052160680294037
step: 190, loss: 0.05293404310941696
step: 200, loss: 0.03804341331124306
step: 210, loss: 0.019524386152625084
/usr/local/lib/python3.7/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=0.5330396475770925, f1=0.508235294117647,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005460625397972763
step: 10, loss: 0.0008718427270650864
step: 20, loss: 0.037672944366931915
step: 30, loss: 0.0004456651513464749
step: 40, loss: 0.0015423489967361093
step: 50, loss: 0.00042641855543479323
step: 60, loss: 0.021923813968896866
step: 70, loss: 0.0036748191341757774
step: 80, loss: 0.0019309561466798186
step: 90, loss: 0.08084122836589813
step: 100, loss: 0.001204988919198513
step: 110, loss: 0.006178372539579868
step: 120, loss: 0.001893025590106845
step: 130, loss: 0.004656339064240456

```

```

step: 140, loss: 0.00031077637686394155
step: 150, loss: 0.0004589996242430061
step: 160, loss: 0.024172117933630943
step: 170, loss: 0.02728012576699257
step: 180, loss: 0.0009756574872881174
step: 190, loss: 0.008068578317761421
step: 200, loss: 0.0026153603103011847
step: 210, loss: 0.003376274136826396
/usr/local/lib/python3.7/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=0.556910569105691, f1=0.543046357615894,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022503137588500977
step: 10, loss: 0.0005302744102664292
step: 20, loss: 0.005348918493837118
step: 30, loss: 0.005420676898211241
step: 40, loss: 0.04753251373767853
step: 50, loss: 0.000923704938031733
step: 60, loss: 0.0008088494068942964
step: 70, loss: 0.0027613218408077955
step: 80, loss: 0.0007949788705445826
step: 90, loss: 0.0005574075621552765
step: 100, loss: 0.007581193931400776
step: 110, loss: 0.007024209015071392
step: 120, loss: 0.00024529866641387343
step: 130, loss: 9.832075738813728e-05
step: 140, loss: 0.00015653362788725644
step: 150, loss: 0.00021725051919929683
step: 160, loss: 0.0011821076041087508
step: 170, loss: 0.0006348742172122002
step: 180, loss: 0.00040372100193053484
step: 190, loss: 0.0005132234655320644
step: 200, loss: 0.0012270512524992228
step: 210, loss: 0.06475219875574112
/usr/local/lib/python3.7/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=0.5107398568019093, f1=0.4742268041237113,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05542132258415222
step: 10, loss: 8.469398017041385e-05

```



```

step: 20, loss: 0.0069498829543590546
step: 30, loss: 0.00959437619894743
step: 40, loss: 0.0018173415446653962
step: 50, loss: 0.008511856198310852
step: 60, loss: 0.0010138709330931306
step: 70, loss: 0.017801614478230476
step: 80, loss: 0.0012714433250948787
step: 90, loss: 0.0007039040210656822
step: 100, loss: 0.0006800146657042205
step: 110, loss: 0.0002499973925296217
step: 120, loss: 0.000603976019192487
step: 130, loss: 0.001081450143828988
step: 140, loss: 0.0003570280096028
step: 150, loss: 0.0007079396164044738
step: 160, loss: 0.00026186660397797823
step: 170, loss: 0.00045459737884812057
step: 180, loss: 0.013819343410432339
step: 190, loss: 0.00011978026304859668
step: 200, loss: 0.0016958791529759765
step: 210, loss: 0.003698096377775073
/usr/local/lib/python3.7/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=0.5242290748898678, f1=0.5165876777251185,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000257878506090492
step: 10, loss: 0.03282354399561882
step: 20, loss: 0.003046554047614336
step: 30, loss: 0.009956438094377518
step: 40, loss: 0.0006671743467450142
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.1081707701086998
step: 60, loss: 0.0006979054887779057
step: 70, loss: 0.004401280079036951
step: 80, loss: 0.0007578208460472524
step: 90, loss: 0.000140794349135831
step: 100, loss: 0.002218192908912897
step: 110, loss: 0.00029516516951844096
step: 120, loss: 0.002163197845220566
step: 130, loss: 0.00039566855411976576
step: 140, loss: 0.008693084120750427
step: 150, loss: 0.00028418810688890517
step: 160, loss: 0.0006778179667890072
step: 170, loss: 0.0001628634927328676
step: 180, loss: 0.0011600106954574585

```

```

step: 190, loss: 0.0005682278424501419
step: 200, loss: 0.001304534380324185
step: 210, loss: 0.00022235827054828405
/usr/local/lib/python3.7/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=0.5277161862527716, f1=0.5190476190476191,
best_f1=0.5934959349593495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015229473821818829
step: 10, loss: 0.00025490421103313565
step: 20, loss: 0.00020496369688771665
step: 30, loss: 0.010550722479820251
step: 40, loss: 0.0001028102487907745
step: 50, loss: 0.00029097325750626624
step: 60, loss: 0.001087950891815126
step: 70, loss: 0.004389083944261074
step: 80, loss: 0.0006607783725485206
step: 90, loss: 0.00022648985031992197
step: 100, loss: 0.0002824841358233243
step: 110, loss: 9.691392915556207e-05
step: 120, loss: 0.0006222878582775593
step: 130, loss: 0.001195160555653274
step: 140, loss: 8.779262861935422e-05
step: 150, loss: 0.00012755939678754658
step: 160, loss: 0.00014791113790124655
step: 170, loss: 0.00012034681276418269
step: 180, loss: 0.00014723972708452493
step: 190, loss: 0.00011996447574347258
step: 200, loss: 0.0002189900551456958
step: 210, loss: 0.0002800129004754126
/usr/local/lib/python3.7/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=0.5249999999999999, f1=0.464, best_f1=0.5934959349593495
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:05, 423.73it/s]
load_f1 = 0.5798816568047337
real_f1 = 0.5810276679841896
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 243.57it/s]

```

3.1.5 Walmart-Amazon - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more

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

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

```
step: 0, loss: 0.5555002689361572
step: 10, loss: 0.38304609060287476
step: 20, loss: 0.30209189653396606
step: 30, loss: 0.43779221177101135
step: 40, loss: 0.42371729016304016
step: 50, loss: 0.2944827973842621
step: 60, loss: 0.278381884098053
step: 70, loss: 0.28037798404693604
step: 80, loss: 0.2355736494064331
step: 90, loss: 0.31915152072906494
step: 100, loss: 0.2926880121231079
step: 110, loss: 0.4175347089767456
step: 120, loss: 0.218088760972023
step: 130, loss: 0.1525937020778656
step: 140, loss: 0.10735838115215302
step: 150, loss: 0.20013992488384247
step: 160, loss: 0.15285204350948334
step: 170, loss: 0.2977769374847412
step: 180, loss: 0.023719223216176033
step: 190, loss: 0.31075096130371094
```

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

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.646900269541779, f1=0.6505376344086021,
best_f1=0.6505376344086021
```

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

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22119615972042084
step: 10, loss: 0.2043842226266861
step: 20, loss: 0.1050400361418724
step: 30, loss: 0.21100960671901703
step: 40, loss: 0.12787044048309326
step: 50, loss: 0.054559119045734406
step: 60, loss: 0.24817313253879547
step: 70, loss: 0.18798361718654633
step: 80, loss: 0.1328384429216385
step: 90, loss: 0.21246208250522614
step: 100, loss: 0.06575346738100052
step: 110, loss: 0.14683039486408234
```

```

step: 120, loss: 0.2812959849834442
step: 130, loss: 0.22164784371852875
step: 140, loss: 0.03856353834271431
step: 150, loss: 0.1597195565700531
step: 160, loss: 0.06644729524850845
step: 170, loss: 0.17604249715805054
step: 180, loss: 0.1411619335412979
step: 190, loss: 0.1004469022154808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7621776504297996, f1=0.7556818181818182,
best_f1=0.7556818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.155671164393425
step: 10, loss: 0.06969989836215973
step: 20, loss: 0.057928960770368576
step: 30, loss: 0.034388985484838486
step: 40, loss: 0.12282807379961014
step: 50, loss: 0.2685750722885132
step: 60, loss: 0.08492142707109451
step: 70, loss: 0.0999271348118782
step: 80, loss: 0.11870589852333069
step: 90, loss: 0.0502941831946373
step: 100, loss: 0.03443985804915428
step: 110, loss: 0.027922378852963448
step: 120, loss: 0.04218433052301407
step: 130, loss: 0.015147285535931587
step: 140, loss: 0.0775231122970581
step: 150, loss: 0.10133138298988342
step: 160, loss: 0.027730504050850868
step: 170, loss: 0.23697338998317719
step: 180, loss: 0.034870680421590805
step: 190, loss: 0.06641225516796112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7875354107648725, f1=0.7867036011080332,

```

```

best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026787763461470604
step: 10, loss: 0.1456955224275589
step: 20, loss: 0.15293757617473602
step: 30, loss: 0.03373951092362404
step: 40, loss: 0.11419711261987686
step: 50, loss: 0.1936606466770172
step: 60, loss: 0.07979127764701843
step: 70, loss: 0.07623796164989471
step: 80, loss: 0.03598848730325699
step: 90, loss: 0.01404961384832859
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.048261139541864395
step: 110, loss: 0.021832125261425972
step: 120, loss: 0.060904838144779205
step: 130, loss: 0.5256698131561279
step: 140, loss: 0.031488481909036636
step: 150, loss: 0.01968127302825451
step: 160, loss: 0.013188500888645649
step: 170, loss: 0.09671128541231155
step: 180, loss: 0.010913411155343056
step: 190, loss: 0.10445568710565567
/usr/local/lib/python3.7/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=0.7639257294429709, f1=0.7454068241469817,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10911764204502106
step: 10, loss: 0.006884886883199215
step: 20, loss: 0.0440412238240242
step: 30, loss: 0.09248389303684235
step: 40, loss: 0.020974138751626015
step: 50, loss: 0.11486558616161346
step: 60, loss: 0.07142658531665802
step: 70, loss: 0.025564461946487427
step: 80, loss: 0.0028468742966651917
step: 90, loss: 0.002501310082152486
step: 100, loss: 0.005938231013715267
step: 110, loss: 0.05073457211256027
step: 120, loss: 0.01192544400691986
step: 130, loss: 0.04125773534178734
step: 140, loss: 0.0959438681602478

```

```

step: 150, loss: 0.17452844977378845
step: 160, loss: 0.06228426471352577
step: 170, loss: 0.03049531765282154
step: 180, loss: 0.12652334570884705
step: 190, loss: 0.11744318157434464
/usr/local/lib/python3.7/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=0.7814207650273224, f1=0.7816711590296496,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008777360431849957
step: 10, loss: 0.07352150976657867
step: 20, loss: 0.0026239294093102217
step: 30, loss: 0.0022848567459732294
step: 40, loss: 0.004713581874966621
step: 50, loss: 0.008849569596350193
step: 60, loss: 0.09317385405302048
step: 70, loss: 0.006336301565170288
step: 80, loss: 0.039267636835575104
step: 90, loss: 0.004189926665276289
step: 100, loss: 0.0005169293144717813
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.0012497773859649897
step: 120, loss: 0.04808416962623596
step: 130, loss: 0.004876546561717987
step: 140, loss: 0.06154170632362366
step: 150, loss: 0.04863117262721062
step: 160, loss: 0.02255800925195217
step: 170, loss: 0.09178189188241959
step: 180, loss: 0.06475553661584854
step: 190, loss: 0.04309706762433052
/usr/local/lib/python3.7/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=0.7357512953367874, f1=0.7216494845360825,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018924265168607235
step: 10, loss: 0.025133337825536728
step: 20, loss: 0.004661178681999445
step: 30, loss: 0.21430878341197968
step: 40, loss: 0.00559197599068284
step: 50, loss: 0.0009666364639997482

```



```

step: 60, loss: 0.04916699230670929
step: 70, loss: 0.0049733989872038364
step: 80, loss: 0.14206786453723907
step: 90, loss: 0.0007175361388362944
step: 100, loss: 0.003778639016672969
step: 110, loss: 0.018581967800855637
step: 120, loss: 0.0007770330994389951
step: 130, loss: 0.0058821141719818115
step: 140, loss: 0.0006813856307417154
step: 150, loss: 0.020756494253873825
step: 160, loss: 0.15633037686347961
step: 170, loss: 0.03776625171303749
step: 180, loss: 0.0012516339775174856
step: 190, loss: 0.0023630510549992323
/usr/local/lib/python3.7/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=0.7719298245614035, f1=0.7443037974683544,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014782330021262169
step: 10, loss: 0.006219899281859398
step: 20, loss: 0.002572414930909872
step: 30, loss: 0.019325172528624535
step: 40, loss: 0.008041621185839176
step: 50, loss: 0.002668553264811635
step: 60, loss: 0.013567655347287655
step: 70, loss: 0.00122905895113945
step: 80, loss: 9.49538080021739e-05
step: 90, loss: 0.00018222095968667418
step: 100, loss: 0.05583369731903076
step: 110, loss: 0.00042024889262393117
step: 120, loss: 0.026221590116620064
step: 130, loss: 0.002496815752238035
step: 140, loss: 0.019143128767609596
step: 150, loss: 0.008364981040358543
step: 160, loss: 0.00042608301737345755
step: 170, loss: 0.0009413987281732261
step: 180, loss: 0.12450702488422394
step: 190, loss: 0.0028828189242631197
/usr/local/lib/python3.7/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=0.7506053268765133, f1=0.7654320987654321,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03874809294939041
step: 10, loss: 0.0167844258248806
step: 20, loss: 0.0003834537055809051
step: 30, loss: 0.00045553912059403956
step: 40, loss: 0.0001770957896951586
step: 50, loss: 0.00032196601387113333
step: 60, loss: 0.00013510597636923194
step: 70, loss: 0.0002754188026301563
step: 80, loss: 0.010244962759315968
step: 90, loss: 0.010241245850920677
step: 100, loss: 0.024476367980241776
step: 110, loss: 0.00038550785393454134
step: 120, loss: 0.0008008843287825584
step: 130, loss: 0.018221894279122353
step: 140, loss: 0.001158858765847981
step: 150, loss: 0.0019997067283838987
step: 160, loss: 0.0002280019543832168
step: 170, loss: 0.0005352155421860516
step: 180, loss: 0.0006096386932767928
step: 190, loss: 0.0005738151376135647
/usr/local/lib/python3.7/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=0.7653061224489796, f1=0.7692307692307693,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015534216072410345
step: 10, loss: 0.00031861240859143436
step: 20, loss: 0.0007227195892482996
step: 30, loss: 0.00045617952127940953
step: 40, loss: 0.0012911271769553423
step: 50, loss: 0.0025922958739101887
step: 60, loss: 0.05567977577447891
step: 70, loss: 0.0024303775280714035
step: 80, loss: 0.00743902288377285
step: 90, loss: 0.002990977605804801
step: 100, loss: 0.0006981004844419658
step: 110, loss: 0.006259371992200613
step: 120, loss: 0.020874345675110817
step: 130, loss: 0.0011734779691323638
step: 140, loss: 0.0058084107004106045
step: 150, loss: 0.00069885776611045
step: 160, loss: 0.000703156110830605
step: 170, loss: 0.000456080335425213

```

```

step: 180, loss: 0.0002152710221707821
step: 190, loss: 0.0004926270339637995
/usr/local/lib/python3.7/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=0.7671957671957672, f1=0.7777777777777779,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000531922560185194
step: 10, loss: 0.08504462987184525
step: 20, loss: 0.0009533062693662941
step: 30, loss: 0.007552796974778175
step: 40, loss: 0.00022269869805313647
step: 50, loss: 0.001643780036829412
step: 60, loss: 0.002408891450613737
step: 70, loss: 0.006673370487987995
step: 80, loss: 0.0003937208093702793
step: 90, loss: 0.002150784945115447
step: 100, loss: 0.009599050506949425
step: 110, loss: 0.00013309833593666553
step: 120, loss: 0.00040164703386835754
step: 130, loss: 0.00012780931137967855
step: 140, loss: 0.0004931005532853305
step: 150, loss: 0.00017819899949245155
step: 160, loss: 0.00017988057516049594
step: 170, loss: 0.0035202682483941317
step: 180, loss: 8.41444925754331e-05
step: 190, loss: 0.00018525564519222826
/usr/local/lib/python3.7/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=0.7762803234501348, f1=0.7642276422764227,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.95456871856004e-05
step: 10, loss: 0.013993223197758198
step: 20, loss: 0.010929716750979424
step: 30, loss: 0.10285363346338272
step: 40, loss: 0.0001270111242774874
step: 50, loss: 0.008706922642886639
step: 60, loss: 0.0007979821530170739
step: 70, loss: 0.00011129546328447759
step: 80, loss: 0.0013382644392549992
step: 90, loss: 0.0003973522107116878

```

```

step: 100, loss: 0.0002502251591067761
step: 110, loss: 0.0011050984030589461
step: 120, loss: 0.0002400847733952105
step: 130, loss: 0.00014663016190752387
step: 140, loss: 0.0004246676980983466
step: 150, loss: 0.01565268263220787
step: 160, loss: 5.296391827869229e-05
step: 170, loss: 0.0344000905752182
step: 180, loss: 0.000316663587000221
step: 190, loss: 0.00010378553270129487
/usr/local/lib/python3.7/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=0.7802197802197802, f1=0.779291553133515,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011208272771909833
step: 10, loss: 0.0001349608355667442
step: 20, loss: 0.0009018585551530123
step: 30, loss: 0.0004931285511702299
step: 40, loss: 0.00037308456376194954
step: 50, loss: 0.0004983377875760198
step: 60, loss: 0.00020628562197089195
step: 70, loss: 0.0019028045935556293
step: 80, loss: 0.00013362016761675477
step: 90, loss: 0.009054397232830524
step: 100, loss: 0.0017012344906106591
step: 110, loss: 0.00027605201466940343
step: 120, loss: 0.0003597010509110987
step: 130, loss: 7.861099584260955e-05
step: 140, loss: 0.0001481144136050716
step: 150, loss: 9.254430915461853e-05
step: 160, loss: 0.00015465958858840168
step: 170, loss: 0.0002126759645761922
step: 180, loss: 5.1846709538949654e-05
step: 190, loss: 0.0003216741606593132
/usr/local/lib/python3.7/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=0.7794871794871795, f1=0.7591623036649214,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022129637363832444
step: 10, loss: 9.853117808233947e-05

```

```

step: 20, loss: 0.000505182018969208
step: 30, loss: 8.01245478214696e-05
step: 40, loss: 0.00022342860756907612
step: 50, loss: 0.0005128324264660478
step: 60, loss: 0.0006875399267300963
step: 70, loss: 0.008746804669499397
step: 80, loss: 0.003677128814160824
step: 90, loss: 0.00260955560952425
step: 100, loss: 0.0006242469535209239
step: 110, loss: 0.00020027671416755766
step: 120, loss: 0.0004471265710890293
step: 130, loss: 0.0003659696376416832
step: 140, loss: 0.00032285842462442815
step: 150, loss: 0.0010223857825621963
step: 160, loss: 0.0005674650892615318
step: 170, loss: 0.00012699908984359354
step: 180, loss: 0.0002480549446772784
step: 190, loss: 0.00015256779443006963
/usr/local/lib/python3.7/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=0.7771739130434783, f1=0.7603305785123966,
best_f1=0.7867036011080332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003561663907021284
step: 10, loss: 9.760217653820291e-05
step: 20, loss: 0.03775252029299736
step: 30, loss: 0.00015340009122155607
step: 40, loss: 0.00012873839295934886
step: 50, loss: 0.00018976478895638138
step: 60, loss: 9.901171870296821e-05
step: 70, loss: 0.0024051503278315067
step: 80, loss: 0.0020226251799613237
step: 90, loss: 0.0009387926547788084
step: 100, loss: 0.0004633224743884057
step: 110, loss: 6.073417534935288e-05
step: 120, loss: 0.00041160083492286503
step: 130, loss: 0.00015963434998411685
step: 140, loss: 0.0002468133461661637
step: 150, loss: 0.0005094018415547907
step: 160, loss: 0.00015724058903288096
step: 170, loss: 0.00019068270921707153
step: 180, loss: 0.0010709903435781598
step: 190, loss: 0.00023996546224225312
/usr/local/lib/python3.7/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=0.7826086956521738, f1=0.7704918032786885,
best_f1=0.7867036011080332
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, 208.99it/s]
load_f1 = 0.7197802197802198
real_f1 = 0.7093333333333335
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:19, 230.50it/s]

```

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6463596820831299
step: 10, loss: 0.37081456184387207
step: 20, loss: 0.29349222779273987
step: 30, loss: 0.38036930561065674
step: 40, loss: 0.28605371713638306
step: 50, loss: 0.23453526198863983
step: 60, loss: 0.2897190451622009
step: 70, loss: 0.3932710289955139
step: 80, loss: 0.3713228404521942
step: 90, loss: 0.23983456194400787
step: 100, loss: 0.13550317287445068
step: 110, loss: 0.28444957733154297
step: 120, loss: 0.18514621257781982
step: 130, loss: 0.049292802810668945
step: 140, loss: 0.17334915697574615
step: 150, loss: 0.2655653953552246
step: 160, loss: 0.153289794921875
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.1742313653230667
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 1: dev_f1=0.7004608294930875, f1=0.7053364269141531,
best_f1=0.7053364269141531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18248677253723145
step: 10, loss: 0.14792653918266296
step: 20, loss: 0.05869488790631294
step: 30, loss: 0.21271076798439026
step: 40, loss: 0.08707400411367416
step: 50, loss: 0.2585260570049286
step: 60, loss: 0.12115494161844254
step: 70, loss: 0.07282178103923798
step: 80, loss: 0.13298864662647247
step: 90, loss: 0.0783933624625206
step: 100, loss: 0.11735516786575317
step: 110, loss: 0.0709221139550209
step: 120, loss: 0.10868120193481445
step: 130, loss: 0.06314220279455185
step: 140, loss: 0.24234509468078613
step: 150, loss: 0.26451805233955383
step: 160, loss: 0.10994374006986618
step: 170, loss: 0.03799218684434891
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7342342342342342, f1=0.7218683651804672,
best_f1=0.7218683651804672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08946157246828079
step: 10, loss: 0.07770045846700668
step: 20, loss: 0.04789634793996811
step: 30, loss: 0.1925671249628067
step: 40, loss: 0.09153319150209427
step: 50, loss: 0.1223522424697876
step: 60, loss: 0.06492840498685837
step: 70, loss: 0.16722311079502106
step: 80, loss: 0.05822918936610222

```

```

step: 90, loss: 0.10986433923244476
step: 100, loss: 0.014474479481577873
step: 110, loss: 0.0824284479022026
step: 120, loss: 0.056185875087976456
step: 130, loss: 0.1629621833562851
step: 140, loss: 0.061491016298532486
step: 150, loss: 0.01743566058576107
step: 160, loss: 0.04860079661011696
step: 170, loss: 0.043738070875406265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7741935483870968, f1=0.7710843373493976,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02032398246228695
step: 10, loss: 0.05956144258379936
step: 20, loss: 0.005683074239641428
step: 30, loss: 0.1055532917380333
step: 40, loss: 0.0030564111657440662
step: 50, loss: 0.010131241753697395
step: 60, loss: 0.11828386038541794
step: 70, loss: 0.07883215695619583
step: 80, loss: 0.05921477451920509
step: 90, loss: 0.14972026646137238
step: 100, loss: 0.2636762857437134
step: 110, loss: 0.08913785964250565
step: 120, loss: 0.01226942241191864
step: 130, loss: 0.1201789528131485
step: 140, loss: 0.04547909274697304
step: 150, loss: 0.07831522077322006
step: 160, loss: 0.13326723873615265
step: 170, loss: 0.0038885045796632767
/usr/local/lib/python3.7/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=0.7409090909090909, f1=0.776824034334764,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0074643283151090145

```

```

step: 10, loss: 0.013701575808227062
step: 20, loss: 0.007994293235242367
step: 30, loss: 0.17220038175582886
step: 40, loss: 0.06216046214103699
step: 50, loss: 0.06591878086328506
step: 60, loss: 0.022783499211072922
step: 70, loss: 0.20804505050182343
step: 80, loss: 0.06982137262821198
step: 90, loss: 0.06510666757822037
step: 100, loss: 0.027164600789546967
step: 110, loss: 0.136735200881958
step: 120, loss: 0.01688736490905285
step: 130, loss: 0.007159671746194363
step: 140, loss: 0.007534175645560026
step: 150, loss: 0.13173650205135345
step: 160, loss: 0.03423348441720009
step: 170, loss: 0.012301350012421608
/usr/local/lib/python3.7/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=0.7566371681415928, f1=0.7435897435897436,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017926543951034546
step: 10, loss: 0.008903006091713905
step: 20, loss: 0.024406615644693375
step: 30, loss: 0.08181920647621155
step: 40, loss: 0.05999622121453285
step: 50, loss: 0.1002536341547966
step: 60, loss: 0.019892234355211258
step: 70, loss: 0.059065792709589005
step: 80, loss: 0.029820796102285385
step: 90, loss: 0.0699053481221199
step: 100, loss: 0.006671664770692587
step: 110, loss: 0.004515441600233316
step: 120, loss: 0.029275601729750633
step: 130, loss: 0.012951986864209175
step: 140, loss: 0.014773006550967693
step: 150, loss: 0.05062452331185341
step: 160, loss: 0.10281651467084885
step: 170, loss: 0.016681550070643425
/usr/local/lib/python3.7/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=0.7368421052631579, f1=0.7418655097613883,
best_f1=0.7710843373493976

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006691332440823317
step: 10, loss: 0.02215990051627159
step: 20, loss: 0.004066120833158493
step: 30, loss: 0.007827671244740486
step: 40, loss: 0.005227426066994667
step: 50, loss: 0.014502519741654396
step: 60, loss: 0.0030974471010267735
step: 70, loss: 0.0011849136790260673
step: 80, loss: 0.007791020907461643
step: 90, loss: 0.0003806964377872646
step: 100, loss: 0.0004747809434775263
step: 110, loss: 0.0801863819360733
step: 120, loss: 0.12298990041017532
step: 130, loss: 0.08408375084400177
step: 140, loss: 0.0012473085662350059
step: 150, loss: 0.007076252717524767
step: 160, loss: 0.0006461804732680321
step: 170, loss: 0.11630506813526154
/usr/local/lib/python3.7/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=0.7570093457943925, f1=0.756152125279642,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015870077535510063
step: 10, loss: 0.010992676019668579
step: 20, loss: 0.0007789698429405689
step: 30, loss: 0.008663278073072433
step: 40, loss: 0.0001767999492585659
step: 50, loss: 0.0007124892435967922
step: 60, loss: 0.0005370856379158795
step: 70, loss: 0.036314643919467926
step: 80, loss: 0.012077064253389835
step: 90, loss: 0.006340591236948967
step: 100, loss: 0.012827537022531033
step: 110, loss: 0.05270221829414368
step: 120, loss: 0.0006728018051944673
step: 130, loss: 0.00119297846686095
step: 140, loss: 0.0005578501732088625
step: 150, loss: 0.013906922191381454
step: 160, loss: 0.01580508053302765
step: 170, loss: 0.03356856852769852
/usr/local/lib/python3.7/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=0.7298578199052131, f1=0.7471526195899771,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013303111307322979
step: 10, loss: 0.009921716526150703
step: 20, loss: 0.0005759168416261673
step: 30, loss: 0.012155490927398205
step: 40, loss: 0.018661992624402046
step: 50, loss: 0.0003940667666029185
step: 60, loss: 0.007771521806716919
step: 70, loss: 0.004592138342559338
step: 80, loss: 0.026826513931155205
step: 90, loss: 0.07094484567642212
step: 100, loss: 0.006324670743197203
step: 110, loss: 0.0009465843322686851
step: 120, loss: 0.011756704188883305
step: 130, loss: 0.09920533746480942
step: 140, loss: 0.013718795962631702
step: 150, loss: 0.020964907482266426
step: 160, loss: 0.04293414577841759
step: 170, loss: 0.010934860445559025
/usr/local/lib/python3.7/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=0.7587939698492463, f1=0.7885985748218527,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01719982922077179
step: 10, loss: 0.00028498820029199123
step: 20, loss: 0.0482509545981884
step: 30, loss: 0.005195870995521545
step: 40, loss: 0.0010723742889240384
step: 50, loss: 0.09259033203125
step: 60, loss: 0.0004691394860856235
step: 70, loss: 0.000503784918691963
step: 80, loss: 0.015697432681918144
step: 90, loss: 0.01731240563094616
step: 100, loss: 0.0008042598492465913
step: 110, loss: 0.008731737732887268
step: 120, loss: 0.00039503860170952976
step: 130, loss: 0.00039943252340890467
step: 140, loss: 0.0070729623548686504

```

```

step: 150, loss: 0.007155211642384529
step: 160, loss: 0.00036041700514033437
step: 170, loss: 0.0005572828813455999
/usr/local/lib/python3.7/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=0.7414634146341463, f1=0.7735849056603774,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06454505771398544
step: 10, loss: 0.0038697714917361736
step: 20, loss: 0.008692799136042595
step: 30, loss: 0.00019410545064602047
step: 40, loss: 0.0006010676152072847
step: 50, loss: 0.0012667240807786584
step: 60, loss: 0.00724773108959198
step: 70, loss: 0.00019614998018369079
step: 80, loss: 0.00020462377869989723
step: 90, loss: 0.0006535116699524224
step: 100, loss: 0.003950489219278097
step: 110, loss: 0.0006750618922524154
step: 120, loss: 0.00021996608120389283
step: 130, loss: 9.779600077308714e-05
step: 140, loss: 0.00024925594334490597
step: 150, loss: 0.00045368928113020957
step: 160, loss: 0.005103704985231161
step: 170, loss: 0.00042858373490162194
/usr/local/lib/python3.7/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=0.7307692307692308, f1=0.7777777777777777,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042346559348516166
step: 10, loss: 0.00018327661382500082
step: 20, loss: 0.0009208688279613853
step: 30, loss: 0.00015365683066193014
step: 40, loss: 0.0010833509732037783
step: 50, loss: 0.00022514787269756198
step: 60, loss: 0.060135699808597565
step: 70, loss: 0.05971459299325943
step: 80, loss: 0.00026441941736266017
step: 90, loss: 0.0005610834923572838
step: 100, loss: 0.002465896075591445

```

```

step: 110, loss: 0.0002625125925987959
step: 120, loss: 0.0008542611030861735
step: 130, loss: 0.012167147360742092
step: 140, loss: 0.0019891478586941957
step: 150, loss: 0.005256480071693659
step: 160, loss: 0.001646367134526372
step: 170, loss: 0.00012021266593364999
/usr/local/lib/python3.7/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=0.7223719676549866, f1=0.7786259541984734,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0065681831911206245
step: 10, loss: 0.0001440257328795269
step: 20, loss: 0.008231619372963905
step: 30, loss: 9.049126674653962e-05
step: 40, loss: 0.014156677760183811
step: 50, loss: 0.013943157158792019
step: 60, loss: 0.001947425538673997
step: 70, loss: 0.00016198903904296458
step: 80, loss: 0.0001337893190793693
step: 90, loss: 0.0001075130348908715
step: 100, loss: 0.048291806131601334
step: 110, loss: 0.0002488842001184821
step: 120, loss: 0.023585207760334015
step: 130, loss: 0.00014160695718601346
step: 140, loss: 0.0002132932422682643
step: 150, loss: 0.005098680965602398
step: 160, loss: 0.00037599849747493863
step: 170, loss: 0.00042425389983691275
/usr/local/lib/python3.7/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=0.7220779220779221, f1=0.7673267326732675,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007745378534309566
step: 10, loss: 9.487292118137702e-05
step: 20, loss: 0.00032110291067510843
step: 30, loss: 0.0014885400887578726
step: 40, loss: 0.0008918948005884886
step: 50, loss: 0.0005758758052252233
step: 60, loss: 0.0005754784215241671

```

```

step: 70, loss: 0.013988837599754333
step: 80, loss: 0.1079673171043396
step: 90, loss: 0.00011620145960478112
step: 100, loss: 0.0003863341989926994
step: 110, loss: 0.0001411228731740266
step: 120, loss: 0.023204421624541283
step: 130, loss: 0.009510360658168793
step: 140, loss: 0.016529077664017677
step: 150, loss: 0.0001207666500704363
step: 160, loss: 0.0001513221359346062
step: 170, loss: 0.00013653938367497176
/usr/local/lib/python3.7/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=0.7225130890052356, f1=0.7673267326732675,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011215208360226825
step: 10, loss: 0.0020480207167565823
step: 20, loss: 0.007287169806659222
step: 30, loss: 0.00017839513020589948
step: 40, loss: 0.0037292330525815487
step: 50, loss: 0.00017510208999738097
step: 60, loss: 0.02618837170302868
step: 70, loss: 0.0001018521434161812
step: 80, loss: 0.03275897353887558
step: 90, loss: 0.0001946557022165507
step: 100, loss: 0.0015549651579931378
step: 110, loss: 0.000721589894965291
step: 120, loss: 0.00023020956723485142
step: 130, loss: 0.0013945807004347444
step: 140, loss: 0.00016955233877524734
step: 150, loss: 0.0026638994459062815
step: 160, loss: 0.0008306876407004893
step: 170, loss: 0.00022829203226137906
/usr/local/lib/python3.7/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=0.7222222222222222, f1=0.7614457831325301,
best_f1=0.7710843373493976
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 [00:07, 240.01it/s]
load_f1 = 0.6329113924050633
real_f1 = 0.6129032258064516
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, 192.11it/s]

```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

```

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

```

```

--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

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

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

```

Selected optimization level 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: 0.6070834398269653
step: 10, loss: 0.6141316890716553
step: 20, loss: 0.45300912857055664
step: 30, loss: 0.29580628871917725
step: 40, loss: 0.09271848946809769
step: 50, loss: 0.09960927069187164
step: 60, loss: 0.17624640464782715
step: 70, loss: 0.22878320515155792
step: 80, loss: 0.17911867797374725
step: 90, loss: 0.1144648864865303
step: 100, loss: 0.005224101711064577
step: 110, loss: 0.1519373655319214
step: 120, loss: 0.10358860343694687
step: 130, loss: 0.022601502016186714
step: 140, loss: 0.01091004628688097
step: 150, loss: 0.020196666941046715
step: 160, loss: 0.0037962067872285843
step: 170, loss: 0.1181349828839302
step: 180, loss: 0.15073640644550323
step: 190, loss: 0.11162997037172318
step: 200, loss: 0.14339609444141388
step: 210, loss: 0.004458216950297356
step: 220, loss: 0.01575925387442112
step: 230, loss: 0.07378827035427094
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9653631284916202, f1=0.970917225950783,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017806170508265495

```

```

step: 10, loss: 0.0021993517875671387
step: 20, loss: 0.17781801521778107
step: 30, loss: 0.13056379556655884
step: 40, loss: 0.1322222799062729
step: 50, loss: 0.007833912968635559
step: 60, loss: 0.011484983377158642
step: 70, loss: 0.02325410023331642
step: 80, loss: 0.0069376579485833645
step: 90, loss: 0.07261423021554947
step: 100, loss: 0.055119019001722336
step: 110, loss: 0.13581590354442596
step: 120, loss: 0.07535926252603531
step: 130, loss: 0.023143813014030457
step: 140, loss: 0.00419433880597353
step: 150, loss: 0.005183257628232241
step: 160, loss: 0.04762054607272148
step: 170, loss: 0.009558978490531445
step: 180, loss: 0.016866616904735565
step: 190, loss: 0.006744676735252142
step: 200, loss: 0.00248658936470747
step: 210, loss: 0.00196867436170578
step: 220, loss: 0.18973183631896973
step: 230, loss: 0.0633171945810318
/usr/local/lib/python3.7/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=0.9620535714285715, f1=0.9686800894854586,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006953548174351454
step: 10, loss: 0.034146998077631
step: 20, loss: 0.010649498552083969
step: 30, loss: 0.02773764729499817
step: 40, loss: 0.05316956341266632
step: 50, loss: 0.005285613238811493
step: 60, loss: 0.002657629083842039
step: 70, loss: 0.025931905955076218
step: 80, loss: 0.0010765320621430874
step: 90, loss: 0.039984624832868576
step: 100, loss: 0.0005749971605837345
step: 110, loss: 0.0015047210035845637
step: 120, loss: 0.005902309436351061
step: 130, loss: 0.002618709346279502
step: 140, loss: 0.010858487337827682
step: 150, loss: 0.05837104097008705
step: 160, loss: 0.1517823040485382

```

```

step: 170, loss: 0.01785176806151867
step: 180, loss: 0.006413109600543976
step: 190, loss: 0.01367938332259655
step: 200, loss: 0.007430526427924633
step: 210, loss: 0.005630389787256718
step: 220, loss: 0.029271764680743217
step: 230, loss: 0.0021905263420194387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9712389380530975, f1=0.9690265486725664,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005620024632662535
step: 10, loss: 0.001072219805791974
step: 20, loss: 0.001524211373180151
step: 30, loss: 0.0015284710098057985
step: 40, loss: 0.0015793376369401813
step: 50, loss: 0.00038226169999688864
step: 60, loss: 0.0005604291218332946
step: 70, loss: 0.0005567208281718194
step: 80, loss: 0.0010540001094341278
step: 90, loss: 0.008696317672729492
step: 100, loss: 0.0012350251199677587
step: 110, loss: 0.0011725776130333543
step: 120, loss: 0.017924979329109192
step: 130, loss: 0.0035656001418828964
step: 140, loss: 0.0006734547787345946
step: 150, loss: 0.0818634107708931
step: 160, loss: 0.1252477616071701
step: 170, loss: 0.07422488927841187
step: 180, loss: 0.026814857497811317
step: 190, loss: 0.0007724290480837226
step: 200, loss: 0.07682400196790695
step: 210, loss: 0.011640203185379505
step: 220, loss: 0.0003298709343653172
step: 230, loss: 0.0010659154504537582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9755555555555556, f1=0.970917225950783,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007216429221443832
step: 10, loss: 0.0010551528539508581
step: 20, loss: 0.0009868283523246646
step: 30, loss: 0.00024452406796626747
step: 40, loss: 0.001035051653161645
step: 50, loss: 0.0004460986237972975
step: 60, loss: 0.1374925673007965
step: 70, loss: 0.03134404122829437
step: 80, loss: 0.0006311989855021238
step: 90, loss: 0.00603610510006547
step: 100, loss: 0.0007362778997048736
step: 110, loss: 0.00059182767290622
step: 120, loss: 0.0005488366587087512
step: 130, loss: 0.0012339743552729487
step: 140, loss: 0.00076742545934394
step: 150, loss: 0.00037879880983382463
step: 160, loss: 0.0006557052838616073
step: 170, loss: 0.008212041109800339
step: 180, loss: 0.009684915654361248
step: 190, loss: 0.0902770534157753
step: 200, loss: 0.001519043929874897
step: 210, loss: 0.004727603867650032
step: 220, loss: 0.0013672026107087731
step: 230, loss: 0.003743137698620558
/usr/local/lib/python3.7/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=0.9697648376259798, f1=0.9631284916201117,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05803803727030754
step: 10, loss: 0.0023099551908671856
step: 20, loss: 0.008259539492428303
step: 30, loss: 0.000996050308458507
step: 40, loss: 0.0017961011035367846
step: 50, loss: 0.03207153081893921
step: 60, loss: 0.00020635676628444344
step: 70, loss: 0.0017051745671778917
step: 80, loss: 0.004321695771068335

```

```

step: 90, loss: 0.00036999714211560786
step: 100, loss: 0.005482832435518503
step: 110, loss: 0.0006428771303035319
step: 120, loss: 0.0012644104426726699
step: 130, loss: 0.0003970529942307621
step: 140, loss: 0.0017392027657479048
step: 150, loss: 0.0004643767315428704
step: 160, loss: 0.0002855063648894429
step: 170, loss: 0.0005429546581581235
step: 180, loss: 0.012440003454685211
step: 190, loss: 0.010761274956166744
step: 200, loss: 0.00015838746912777424
step: 210, loss: 0.00016704823065083474
step: 220, loss: 0.0002986571053043008
step: 230, loss: 0.03933864086866379
/usr/local/lib/python3.7/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=0.9722530521642618, f1=0.9645232815964524,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007586805149912834
step: 10, loss: 0.00011166315380251035
step: 20, loss: 0.00012786351726390421
step: 30, loss: 6.786973244743422e-05
step: 40, loss: 0.00026986817829310894
step: 50, loss: 0.00010478062176844105
step: 60, loss: 0.00012280237569939345
step: 70, loss: 0.0069576152600348
step: 80, loss: 0.053543396294116974
step: 90, loss: 0.00026002092636190355
step: 100, loss: 0.0001581356773385778
step: 110, loss: 0.000159014729433693
step: 120, loss: 7.695490785408765e-05
step: 130, loss: 0.00014707326772622764
step: 140, loss: 0.001837821793742478
step: 150, loss: 0.017979197204113007
step: 160, loss: 0.05693701654672623
step: 170, loss: 0.00044129922753199935
step: 180, loss: 0.000440343254012987
step: 190, loss: 0.05220046639442444
step: 200, loss: 0.0034476041328161955
step: 210, loss: 5.779396815341897e-05
step: 220, loss: 0.001088439836166799
step: 230, loss: 0.0001379034947603941
/usr/local/lib/python3.7/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=0.9690949227373068, f1=0.9646799116997793,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.008787360973656e-05
step: 10, loss: 9.520164167042822e-05
step: 20, loss: 0.00011193565296707675
step: 30, loss: 7.784995977999642e-05
step: 40, loss: 0.00038775376742705703
step: 50, loss: 0.0001346348290098831
step: 60, loss: 0.0010771583765745163
step: 70, loss: 0.00015793778584338725
step: 80, loss: 0.0002898851234931499
step: 90, loss: 0.00027002699789591134
step: 100, loss: 0.0002958435798063874
step: 110, loss: 0.00016014916764106601
step: 120, loss: 0.0010759613942354918
step: 130, loss: 8.256528235506266e-05
step: 140, loss: 8.463017729809508e-05
step: 150, loss: 8.322787471115589e-05
step: 160, loss: 0.0002864086418412626
step: 170, loss: 0.00017695198766887188
step: 180, loss: 0.0002473946660757065
step: 190, loss: 0.024896522983908653
step: 200, loss: 0.0010624683927744627
step: 210, loss: 0.0005496293888427317
step: 220, loss: 0.00028160898364149034
step: 230, loss: 0.0005371874431148171
/usr/local/lib/python3.7/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=0.9665178571428571, f1=0.967452300785634,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.7457749790046364e-05
step: 10, loss: 0.0003635529719758779
step: 20, loss: 0.0005586601910181344
step: 30, loss: 0.00019148597493767738
step: 40, loss: 0.000686321291141212
step: 50, loss: 0.00012445247557479888
step: 60, loss: 7.48821476008743e-05
step: 70, loss: 0.00015403417637571692
step: 80, loss: 0.0001646694727241993

```



```
step: 90, loss: 0.00041839637560769916
step: 100, loss: 0.0016333460807800293
step: 110, loss: 3.9143720641732216e-05
step: 120, loss: 4.9166588723892346e-05
step: 130, loss: 0.0001080703004845418
step: 140, loss: 3.978046515840106e-05
step: 150, loss: 0.000177792739123106
step: 160, loss: 5.4887899750610813e-05
step: 170, loss: 5.9441437770146877e-05
step: 180, loss: 0.006337947677820921
step: 190, loss: 0.052125006914138794
step: 200, loss: 0.3091333210468292
step: 210, loss: 0.00023094475909601897
step: 220, loss: 0.00018042924057226628
step: 230, loss: 0.00021629153343383223
/usr/local/lib/python3.7/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=0.9695603156708005, f1=0.9661399548532732,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008055396028794348
step: 10, loss: 0.0005641201860271394
step: 20, loss: 0.00014915312931407243
step: 30, loss: 0.00015654154412914068
step: 40, loss: 0.00014144495071377605
step: 50, loss: 0.0003302981494925916
step: 60, loss: 0.0004331960517447442
step: 70, loss: 0.0009570837137289345
step: 80, loss: 0.00012629773118533194
step: 90, loss: 0.00025552534498274326
step: 100, loss: 8.452798647340387e-05
step: 110, loss: 0.00025974944583140314
step: 120, loss: 0.000108711639768444
step: 130, loss: 0.000689839362166822
step: 140, loss: 0.0024289570283144712
step: 150, loss: 0.00042461484554223716
step: 160, loss: 6.760514224879444e-05
step: 170, loss: 8.708695531822741e-05
step: 180, loss: 0.00013062327343504876
step: 190, loss: 0.001155154430307448
step: 200, loss: 5.3157040383666754e-05
step: 210, loss: 0.0002034911885857582
step: 220, loss: 8.523281576344743e-05
step: 230, loss: 0.00034637501812539995
/usr/local/lib/python3.7/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=0.9710467706013363, f1=0.968609865470852,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002251895348308608
step: 10, loss: 0.00011271484981989488
step: 20, loss: 0.0011299503967165947
step: 30, loss: 0.0016605928540229797
step: 40, loss: 0.00016974593745544553
step: 50, loss: 0.00011215761333005503
step: 60, loss: 0.0008421723614446819
step: 70, loss: 0.00029203793383203447
step: 80, loss: 4.853076097788289e-05
step: 90, loss: 0.00012588083336595446
step: 100, loss: 5.486189911607653e-05
step: 110, loss: 0.0001459281047573313
step: 120, loss: 6.5470980189275e-05
step: 130, loss: 5.575868999585509e-05
step: 140, loss: 4.53682332590688e-05
step: 150, loss: 0.0004478734917938709
step: 160, loss: 4.3013696995330974e-05
step: 170, loss: 9.120254981098697e-05
step: 180, loss: 0.00010618925443850458
step: 190, loss: 3.938256486435421e-05
step: 200, loss: 0.000732608896214515
step: 210, loss: 4.681244172388688e-05
step: 220, loss: 6.116511940490454e-05
step: 230, loss: 6.361435225699097e-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=0.968609865470852, f1=0.972972972972973,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.764043842442334e-05
step: 10, loss: 5.2817260439042e-05
step: 20, loss: 0.0001172860138467513
step: 30, loss: 7.519101200159639e-05
step: 40, loss: 0.00010223700519418344
step: 50, loss: 0.006976041942834854
step: 60, loss: 0.00104851508513093
step: 70, loss: 8.187331695808098e-05
step: 80, loss: 7.396753062494099e-05

```

```

step: 90, loss: 4.4417458411771804e-05
step: 100, loss: 3.8670968933729455e-05
step: 110, loss: 5.174899706616998e-05
step: 120, loss: 5.134172897669487e-05
step: 130, loss: 5.159762076800689e-05
step: 140, loss: 4.319286745158024e-05
step: 150, loss: 0.00017904405831359327
step: 160, loss: 5.4605083278147504e-05
step: 170, loss: 6.260824739001691e-05
step: 180, loss: 6.325754657154903e-05
step: 190, loss: 4.480501593207009e-05
step: 200, loss: 3.068052319576964e-05
step: 210, loss: 4.5096006942912936e-05
step: 220, loss: 5.150603465153836e-05
step: 230, loss: 9.944501653080806e-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=0.9696969696969697, f1=0.9741282339707535,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.7320056182798e-05
step: 10, loss: 5.29545868630521e-05
step: 20, loss: 6.0029149608453736e-05
step: 30, loss: 0.00011591743532335386
step: 40, loss: 0.00010422208288218826
step: 50, loss: 7.737516716588289e-05
step: 60, loss: 4.966248889104463e-05
step: 70, loss: 8.217349386541173e-05
step: 80, loss: 7.507181726396084e-05
step: 90, loss: 0.086146779358387
step: 100, loss: 6.207353726495057e-05
step: 110, loss: 0.02489672414958477
step: 120, loss: 0.0014913877239450812
step: 130, loss: 8.360816718777642e-05
step: 140, loss: 9.189615957438946e-05
step: 150, loss: 4.9242218665312976e-05
step: 160, loss: 3.602624565246515e-05
step: 170, loss: 0.00011888687004102394
step: 180, loss: 0.000262441550148651
step: 190, loss: 6.128395034465939e-05
step: 200, loss: 0.0013151471503078938
step: 210, loss: 4.1557046642992646e-05
step: 220, loss: 7.01846947777085e-05
step: 230, loss: 4.232853098073974e-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=0.9688888888888889, f1=0.9697648376259798,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.787466554669663e-05
step: 10, loss: 0.00022383357281796634
step: 20, loss: 4.1389826947124675e-05
step: 30, loss: 7.311075751204044e-05
step: 40, loss: 0.001000163028948009
step: 50, loss: 0.00038353147101588547
step: 60, loss: 7.426628144457936e-05
step: 70, loss: 5.31412115378771e-05
step: 80, loss: 4.607604569173418e-05
step: 90, loss: 0.00013258328544907272
step: 100, loss: 3.3224834623979405e-05
step: 110, loss: 3.914708941010758e-05
step: 120, loss: 2.4515586119377986e-05
step: 130, loss: 3.16008590743877e-05
step: 140, loss: 4.8519672418478876e-05
step: 150, loss: 3.868892963510007e-05
step: 160, loss: 0.00012911581143271178
step: 170, loss: 2.869496529456228e-05
step: 180, loss: 3.2159579859580845e-05
step: 190, loss: 4.5892807975178584e-05
step: 200, loss: 4.2012048652395606e-05
step: 210, loss: 0.0008986610919237137
step: 220, loss: 9.997133747674525e-05
step: 230, loss: 0.0015923046739771962
/usr/local/lib/python3.7/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=0.9695603156708005, f1=0.975, best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.899678293033503e-05
step: 10, loss: 3.061706229345873e-05
step: 20, loss: 4.3609314161585644e-05
step: 30, loss: 5.007538857171312e-05
step: 40, loss: 6.032371675246395e-05
step: 50, loss: 0.0007473031291738153
step: 60, loss: 5.954835069132969e-05
step: 70, loss: 7.172892219386995e-05
step: 80, loss: 0.008979091420769691
step: 90, loss: 0.00682540750131011

```

```

step: 100, loss: 3.996264058514498e-05
step: 110, loss: 2.936562850663904e-05
step: 120, loss: 5.34495193278417e-05
step: 130, loss: 3.9053469663485885e-05
step: 140, loss: 5.927605525357649e-05
step: 150, loss: 5.617280476144515e-05
step: 160, loss: 2.7964833861915395e-05
step: 170, loss: 2.868370575015433e-05
step: 180, loss: 6.935273995622993e-05
step: 190, loss: 0.00031777011463418603
step: 200, loss: 8.889767195796594e-05
step: 210, loss: 7.807690417394042e-05
step: 220, loss: 7.565190026070923e-05
step: 230, loss: 4.2220199247822165e-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=0.9695603156708005, f1=0.9727891156462585,
best_f1=0.970917225950783
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

    warnings.warn("An input tensor was not cuda.")
2473it [00:13, 187.11it/s]
load_f1 = 0.9700332963374029
real_f1 = 0.9690265486725664
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:19, 229.20it/s]

```

3.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

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

```
step: 0, loss: 0.6171518564224243
step: 10, loss: 0.5168858766555786
step: 20, loss: 0.5265028476715088
step: 30, loss: 0.2412852644920349
step: 40, loss: 0.23249362409114838
step: 50, loss: 0.16938233375549316
step: 60, loss: 0.08005590736865997
step: 70, loss: 0.1135798990726471
step: 80, loss: 0.07488192617893219
step: 90, loss: 0.2645295560359955
step: 100, loss: 0.031099442392587662
step: 110, loss: 0.22898361086845398
step: 120, loss: 0.17071714997291565
step: 130, loss: 0.039370737969875336
step: 140, loss: 0.053496990352869034
```

```

step: 150, loss: 0.08306758105754852
step: 160, loss: 0.05495429411530495
step: 170, loss: 0.1345777064561844
step: 180, loss: 0.05474938824772835
step: 190, loss: 0.034868206828832626
step: 200, loss: 0.21665257215499878
step: 210, loss: 0.11267144978046417
step: 220, loss: 0.2575948238372803
step: 230, loss: 0.17139871418476105
step: 240, loss: 0.10591217875480652
step: 250, loss: 0.02327829599380493
step: 260, loss: 0.2685248851776123
step: 270, loss: 0.020945671945810318
step: 280, loss: 0.07853183150291443
step: 290, loss: 0.059843871742486954
step: 300, loss: 0.03644246608018875
step: 310, loss: 0.2111869752407074
step: 320, loss: 0.07625424116849899
step: 330, loss: 0.019213803112506866
step: 340, loss: 0.08048802614212036
step: 350, loss: 0.03841885179281235
step: 360, loss: 0.12639640271663666
step: 370, loss: 0.19996137917041779
step: 380, loss: 0.016337184235453606
step: 390, loss: 0.18663115799427032
step: 400, loss: 0.3055819272994995
step: 410, loss: 0.09593809396028519
step: 420, loss: 0.05984706059098244
step: 430, loss: 0.16703490912914276
step: 440, loss: 0.031795796006917953
step: 450, loss: 0.013465615920722485
step: 460, loss: 0.012115530669689178
step: 470, loss: 0.104311004281044
step: 480, loss: 0.12388470768928528
step: 490, loss: 0.11551483720541
step: 500, loss: 0.11879421770572662
step: 510, loss: 0.07189424335956573
step: 520, loss: 0.01478670910000801
step: 530, loss: 0.015965767204761505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9182156133828996,
best_f1=0.9182156133828996

```



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

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

```
step: 0, loss: 0.173036128282547  
step: 10, loss: 0.03802628442645073  
step: 20, loss: 0.044142819941043854  
step: 30, loss: 0.018441468477249146  
step: 40, loss: 0.03524432331323624  
step: 50, loss: 0.21605700254440308  
step: 60, loss: 0.021920522674918175  
step: 70, loss: 0.06950458884239197  
step: 80, loss: 0.10971830040216446  
step: 90, loss: 0.0232094656676054  
step: 100, loss: 0.024179281666874886  
step: 110, loss: 0.024556251242756844  
step: 120, loss: 0.057501740753650665  
step: 130, loss: 0.22697964310646057  
step: 140, loss: 0.028336964547634125  
step: 150, loss: 0.061127129942178726  
step: 160, loss: 0.021154671907424927  
step: 170, loss: 0.01783970184624195  
step: 180, loss: 0.04179883003234863  
step: 190, loss: 0.006654696073383093  
step: 200, loss: 0.008101933635771275  
step: 210, loss: 0.04120688885450363  
step: 220, loss: 0.046168845146894455  
step: 230, loss: 0.01981900818645954  
step: 240, loss: 0.01383217517286539  
step: 250, loss: 0.04388715326786041  
step: 260, loss: 0.005367661360651255  
step: 270, loss: 0.26459482312202454  
step: 280, loss: 0.09946729987859726  
step: 290, loss: 0.011302192695438862  
step: 300, loss: 0.11926276981830597  
step: 310, loss: 0.024582814425230026  
step: 320, loss: 0.09062347561120987  
step: 330, loss: 0.03051218017935753  
step: 340, loss: 0.08775606006383896  
step: 350, loss: 0.013399711810052395  
step: 360, loss: 0.14555969834327698  
step: 370, loss: 0.15422193706035614  
step: 380, loss: 0.06506285816431046  
step: 390, loss: 0.07450387626886368  
step: 400, loss: 0.10507708042860031  
step: 410, loss: 0.007496484089642763  
step: 420, loss: 0.045647330582141876  
step: 430, loss: 0.03555232286453247  
step: 440, loss: 0.21121206879615784
```

```

step: 450, loss: 0.01650765724480152
step: 460, loss: 0.028354844078421593
step: 470, loss: 0.11501602828502655
step: 480, loss: 0.25837793946266174
step: 490, loss: 0.0344323068857193
step: 500, loss: 0.2953059673309326
step: 510, loss: 0.02606995776295662
step: 520, loss: 0.05704716965556145
step: 530, loss: 0.023655004799365997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9375573921028467, f1=0.9305936073059361,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047255951911211014
step: 10, loss: 0.038712967187166214
step: 20, loss: 0.0386463925242424
step: 30, loss: 0.009129852056503296
step: 40, loss: 0.032854922115802765
step: 50, loss: 0.14950881898403168
step: 60, loss: 0.00919110793620348
step: 70, loss: 0.011790819466114044
step: 80, loss: 0.0032886292319744825
step: 90, loss: 0.008117054589092731
step: 100, loss: 0.15122805535793304
step: 110, loss: 0.0023060422390699387
step: 120, loss: 0.0021247633267194033
step: 130, loss: 0.0032642162404954433
step: 140, loss: 0.03821643441915512
step: 150, loss: 0.047704074531793594
step: 160, loss: 0.015509921126067638
step: 170, loss: 0.031523626297712326
step: 180, loss: 0.012055655010044575
step: 190, loss: 0.010167093016207218
step: 200, loss: 0.018649065867066383
step: 210, loss: 0.094906747341156
step: 220, loss: 0.05606693774461746
step: 230, loss: 0.04089685529470444
step: 240, loss: 0.0045444779098033905
step: 250, loss: 0.06528500467538834
step: 260, loss: 0.026768315583467484

```

```
step: 270, loss: 0.06466194242238998
step: 280, loss: 0.02246764488518238
step: 290, loss: 0.0020299083553254604
step: 300, loss: 0.010742797516286373
step: 310, loss: 0.012118804268538952
step: 320, loss: 0.019480673596262932
step: 330, loss: 0.0021353072952479124
step: 340, loss: 0.015701230615377426
step: 350, loss: 0.002279554260894656
step: 360, loss: 0.030170176178216934
step: 370, loss: 0.007664240896701813
step: 380, loss: 0.11285746097564697
step: 390, loss: 0.0179061871021986
step: 400, loss: 0.048016563057899475
step: 410, loss: 0.004942705389112234
step: 420, loss: 0.11192207783460617
step: 430, loss: 0.04219295084476471
step: 440, loss: 0.01373417116701603
step: 450, loss: 0.08873149752616882
step: 460, loss: 0.11301665008068085
step: 470, loss: 0.1827819049358368
step: 480, loss: 0.01662040501832962
step: 490, loss: 0.021627767011523247
step: 500, loss: 0.02081793174147606
step: 510, loss: 0.007649534847587347
step: 520, loss: 0.16518747806549072
step: 530, loss: 0.0323701947927475
/usr/local/lib/python3.7/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=0.9293119698397737, f1=0.9216417910447762,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004243796691298485
step: 10, loss: 0.011144411750137806
step: 20, loss: 0.0023875548504292965
step: 30, loss: 0.02877119369804859
step: 40, loss: 0.016184108331799507
step: 50, loss: 0.0050880080088973045
step: 60, loss: 0.05291111394762993
step: 70, loss: 0.0028964728116989136
step: 80, loss: 0.018667476251721382
step: 90, loss: 0.058991532772779465
step: 100, loss: 0.006583447568118572
step: 110, loss: 0.061065033078193665
step: 120, loss: 0.0020717671141028404
```

```

step: 130, loss: 0.0065793609246611595
step: 140, loss: 0.0028591903392225504
step: 150, loss: 0.002655907068401575
step: 160, loss: 0.006048892624676228
step: 170, loss: 0.025693489238619804
step: 180, loss: 0.002879670588299632
step: 190, loss: 0.15577706694602966
step: 200, loss: 0.005636070854961872
step: 210, loss: 0.0967487245798111
step: 220, loss: 0.01334457378834486
step: 230, loss: 0.06920173764228821
step: 240, loss: 0.017569396644830704
step: 250, loss: 0.02333711087703705
step: 260, loss: 0.020459851250052452
step: 270, loss: 0.013945107348263264
step: 280, loss: 0.04587923362851143
step: 290, loss: 0.017359590157866478
step: 300, loss: 0.0011036478681489825
step: 310, loss: 0.06714034825563431
step: 320, loss: 0.0027080136351287365
step: 330, loss: 0.022895516827702522
step: 340, loss: 0.003878965973854065
step: 350, loss: 0.002907809801399708
step: 360, loss: 0.03378571942448616
step: 370, loss: 0.0023232901003211737
step: 380, loss: 0.002385127590969205
step: 390, loss: 0.0020231897942721844
step: 400, loss: 0.006580818444490433
step: 410, loss: 0.06393841654062271
step: 420, loss: 0.0444924496114254
step: 430, loss: 0.08809266239404678
step: 440, loss: 0.005138610024005175
step: 450, loss: 0.0029195942915976048
step: 460, loss: 0.0014144186861813068
step: 470, loss: 0.01345062255859375
step: 480, loss: 0.0037173789460211992
step: 490, loss: 0.008433017879724503
step: 500, loss: 0.008318416774272919
step: 510, loss: 0.05813133344054222
step: 520, loss: 0.029929382726550102
step: 530, loss: 0.01935247890651226
/usr/local/lib/python3.7/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=0.9311778290993071, f1=0.917506874427131,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12825383245944977
step: 10, loss: 0.025657057762145996
step: 20, loss: 0.0029347159434109926
step: 30, loss: 0.023085132241249084
step: 40, loss: 0.0020560957491397858
step: 50, loss: 0.038974661380052567
step: 60, loss: 0.032694678753614426
step: 70, loss: 0.02195911295711994
step: 80, loss: 0.0007001644116826355
step: 90, loss: 0.011352481320500374
step: 100, loss: 0.0022848902735859156
step: 110, loss: 0.014804142527282238
step: 120, loss: 0.007898861542344093
step: 130, loss: 0.03862179070711136
step: 140, loss: 0.006188669707626104
step: 150, loss: 0.005184897221624851
step: 160, loss: 0.00027908963966183364
step: 170, loss: 0.010978414677083492
step: 180, loss: 0.0004936343757435679
step: 190, loss: 0.0008882739348337054
step: 200, loss: 0.004001426510512829
step: 210, loss: 0.011133910156786442
step: 220, loss: 0.0009879147401079535
step: 230, loss: 0.01450782548636198
step: 240, loss: 0.0007690875208936632
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 250, loss: 0.04968202859163284
step: 260, loss: 0.0012776304502040148
step: 270, loss: 0.0005586500628851354
step: 280, loss: 0.0022408622317016125
step: 290, loss: 0.0973658338189125
step: 300, loss: 0.0012497804127633572
step: 310, loss: 0.004847160540521145
step: 320, loss: 0.23541951179504395
step: 330, loss: 0.01414217334240675
step: 340, loss: 0.00367065891623497
step: 350, loss: 0.006504901684820652
step: 360, loss: 0.022393563762307167
step: 370, loss: 0.0004248045152053237
step: 380, loss: 0.0005124150775372982
step: 390, loss: 0.0002912726777140051
step: 400, loss: 0.005603572819381952
step: 410, loss: 0.016436733305454254
step: 420, loss: 0.0013563764514401555
step: 430, loss: 0.008046119473874569
step: 440, loss: 0.001524827559478581
step: 450, loss: 0.03364356979727745
```

```

step: 460, loss: 0.003653785213828087
step: 470, loss: 0.011239934712648392
step: 480, loss: 0.0011384435929358006
step: 490, loss: 0.0005101162241771817
step: 500, loss: 0.0024071987718343735
step: 510, loss: 0.0016554780304431915
step: 520, loss: 0.0057458127848804
step: 530, loss: 0.003770777489989996
/usr/local/lib/python3.7/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=0.9268518518518517, f1=0.9237132352941178,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0504474975168705
step: 10, loss: 0.0007179527310654521
step: 20, loss: 0.0684642344713211
step: 30, loss: 0.0010298504494130611
step: 40, loss: 0.0015070834197103977
step: 50, loss: 0.0480663999915123
step: 60, loss: 0.00015810944023542106
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.0004843227507080883
step: 80, loss: 0.013410845771431923
step: 90, loss: 0.0005134496605023742
step: 100, loss: 0.0019356623524799943
step: 110, loss: 0.02490430884063244
step: 120, loss: 0.0007272568764165044
step: 130, loss: 0.002705308608710766
step: 140, loss: 0.003847420448437333
step: 150, loss: 0.0001893879525596276
step: 160, loss: 0.0003591121349018067
step: 170, loss: 0.009674015454947948
step: 180, loss: 0.00030202927882783115
step: 190, loss: 0.018473705276846886
step: 200, loss: 0.010050513781607151
step: 210, loss: 0.00466138543561101
step: 220, loss: 0.007149404380470514
step: 230, loss: 0.0003297666844446212
step: 240, loss: 0.0036716910544782877
step: 250, loss: 0.018440207466483116
step: 260, loss: 0.0002289334370288998
step: 270, loss: 0.12770673632621765
step: 280, loss: 0.0007068304112181067
step: 290, loss: 0.00015457862173207104
step: 300, loss: 0.025252103805541992

```

```

step: 310, loss: 0.0003365633310750127
step: 320, loss: 0.0004708452324848622
step: 330, loss: 0.0011616195552051067
step: 340, loss: 0.08157515525817871
step: 350, loss: 0.017435265704989433
step: 360, loss: 0.004799480549991131
step: 370, loss: 0.0011369542917236686
step: 380, loss: 0.000161726915393956
step: 390, loss: 0.010509274899959564
step: 400, loss: 0.0050035640597343445
step: 410, loss: 0.0035370546393096447
step: 420, loss: 0.13632631301879883
step: 430, loss: 0.0005424062255769968
step: 440, loss: 0.0004707020416390151
step: 450, loss: 0.00015391931810881943
step: 460, loss: 0.00013699712872039527
step: 470, loss: 0.002894903300330043
step: 480, loss: 0.0013801779132336378
step: 490, loss: 0.00273912469856441
step: 500, loss: 0.001682671019807458
step: 510, loss: 0.008316111750900745
step: 520, loss: 0.0049744597636163235
step: 530, loss: 0.005506511311978102
/usr/local/lib/python3.7/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=0.9281663516068053, f1=0.9213483146067416,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023215626424644142
step: 10, loss: 0.0010444649960845709
step: 20, loss: 0.07201486825942993
step: 30, loss: 0.0002887133159674704
step: 40, loss: 0.004719472024589777
step: 50, loss: 0.0003613114240579307
step: 60, loss: 0.0020716742146760225
step: 70, loss: 0.002629755064845085
step: 80, loss: 0.0003673604514915496
step: 90, loss: 0.0002668345987331122
step: 100, loss: 0.0034270775504410267
step: 110, loss: 0.00045758788473904133
step: 120, loss: 0.0008325062226504087
step: 130, loss: 0.001795317279174924
step: 140, loss: 0.00022444648493546993
step: 150, loss: 0.000857586448546499
step: 160, loss: 0.004196892026811838

```

```

step: 170, loss: 0.002298473846167326
step: 180, loss: 0.0014518167590722442
step: 190, loss: 0.0075037507340312
step: 200, loss: 0.00033034777152352035
step: 210, loss: 0.0006092317053116858
step: 220, loss: 0.0003183274529874325
step: 230, loss: 0.24989330768585205
step: 240, loss: 0.0007684558513574302
step: 250, loss: 0.00689398217946291
step: 260, loss: 0.0012545317877084017
step: 270, loss: 0.0007680127746425569
step: 280, loss: 0.00102240068372339
step: 290, loss: 0.0021741599775850773
step: 300, loss: 0.0023465557023882866
step: 310, loss: 0.0017890315502882004
step: 320, loss: 0.0005237768054939806
step: 330, loss: 0.00023064546985551715
step: 340, loss: 0.08642736822366714
step: 350, loss: 0.0038940091617405415
step: 360, loss: 0.0012391242198646069
step: 370, loss: 0.03565956652164459
step: 380, loss: 0.00040002004243433475
step: 390, loss: 0.000557691091671586
step: 400, loss: 0.02262299694120884
step: 410, loss: 0.0009015344548970461
step: 420, loss: 0.00036853415076620877
step: 430, loss: 9.520693856757134e-05
step: 440, loss: 0.0004806859651580453
step: 450, loss: 0.0003051359672099352
step: 460, loss: 0.00012186478124931455
step: 470, loss: 0.0005103765870444477
step: 480, loss: 0.08846034109592438
step: 490, loss: 0.003367595374584198
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 500, loss: 0.016441820189356804
step: 510, loss: 0.0005467355367727578
step: 520, loss: 0.01242764387279749
step: 530, loss: 0.0006069561350159347
/usr/local/lib/python3.7/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=0.9319664492078286, f1=0.9281105990783408,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002070032060146332
step: 10, loss: 0.04213781654834747

```


step: 20, loss: 0.000745740719139576
step: 30, loss: 0.0014428164577111602
step: 40, loss: 0.0002714419097173959
step: 50, loss: 0.013658042065799236
step: 60, loss: 0.001716212835162878
step: 70, loss: 0.0001799036399461329
step: 80, loss: 0.0005100186099298298
step: 90, loss: 0.0003568484971765429
step: 100, loss: 0.00022701879788655788
step: 110, loss: 0.0002885569410864264
step: 120, loss: 0.00022906278900336474
step: 130, loss: 0.0007178825908340514
step: 140, loss: 0.004146234598010778
step: 150, loss: 0.00013904637307859957
step: 160, loss: 0.0023261273745447397
step: 170, loss: 0.02208457700908184
step: 180, loss: 0.12069432437419891
step: 190, loss: 0.0006274657207541168
step: 200, loss: 0.0058480193838477135
step: 210, loss: 0.0015450933715328574
step: 220, loss: 0.0012390583287924528
step: 230, loss: 0.0011165140895172954
step: 240, loss: 0.1604631543159485
step: 250, loss: 0.004263932351022959
step: 260, loss: 0.00015712293679825962
step: 270, loss: 0.0015262312954291701
step: 280, loss: 0.012826587073504925
step: 290, loss: 0.00027645574300549924
step: 300, loss: 0.0006822083378210664
step: 310, loss: 0.015029375441372395
step: 320, loss: 0.03458826616406441
step: 330, loss: 0.003506455337628722
step: 340, loss: 0.0005167393246665597
step: 350, loss: 0.00019693108333740383
step: 360, loss: 0.01042347215116024
step: 370, loss: 0.00042316486360505223
step: 380, loss: 0.0003682312380988151
step: 390, loss: 0.08180826902389526
step: 400, loss: 0.0011456640204414725
step: 410, loss: 0.00018220093625131994
step: 420, loss: 0.004890593700110912
step: 430, loss: 0.019020909443497658
step: 440, loss: 0.0008893120102584362
step: 450, loss: 0.00017227252828888595
step: 460, loss: 0.0001707534829620272
step: 470, loss: 0.00031927789677865803
step: 480, loss: 0.0001443695364287123
step: 490, loss: 0.06193586066365242

```

step: 500, loss: 0.001947403303347528
step: 510, loss: 0.00393174821510911
step: 520, loss: 0.0011085435980930924
step: 530, loss: 0.00016560746007598937
/usr/local/lib/python3.7/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=0.9252336448598131, f1=0.9190432382704692,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008383772801607847
step: 10, loss: 0.00045192264951765537
step: 20, loss: 0.038867104798555374
step: 30, loss: 0.0005433496553450823
step: 40, loss: 8.781263750279322e-05
step: 50, loss: 0.0013673790963366628
step: 60, loss: 0.00024179606407415122
step: 70, loss: 0.0001289791543968022
step: 80, loss: 0.15514031052589417
step: 90, loss: 0.0036468682810664177
step: 100, loss: 0.0005888910964131355
step: 110, loss: 0.027273274958133698
step: 120, loss: 0.0016963089583441615
step: 130, loss: 0.02651095762848854
step: 140, loss: 0.0015956501010805368
step: 150, loss: 0.00016811367822811007
step: 160, loss: 0.004781712777912617
step: 170, loss: 0.04100522771477699
step: 180, loss: 7.90066915214993e-05
step: 190, loss: 0.00018132355762645602
step: 200, loss: 0.00010275749809807166
step: 210, loss: 0.0017962661804631352
step: 220, loss: 0.0027182360645383596
step: 230, loss: 0.016961006447672844
step: 240, loss: 0.00044860871275886893
step: 250, loss: 0.00013619598757941276
step: 260, loss: 0.004530719481408596
step: 270, loss: 4.820555477635935e-05
step: 280, loss: 0.0003097834705840796
step: 290, loss: 0.04432355612516403
step: 300, loss: 9.092967957258224e-05
step: 310, loss: 0.0037694573402404785
step: 320, loss: 0.00019280653214082122
step: 330, loss: 0.00018937034474220127
step: 340, loss: 0.00014481969992630184
step: 350, loss: 0.0006035887054167688

```

```

step: 360, loss: 6.540115282405168e-05
step: 370, loss: 0.0001035236127790995
step: 380, loss: 0.008500290103256702
step: 390, loss: 7.278526027221233e-05
step: 400, loss: 0.0001544381375424564
step: 410, loss: 0.00010451031994307414
step: 420, loss: 6.484125333372504e-05
step: 430, loss: 0.00017343157378491014
step: 440, loss: 0.0001226544554810971
step: 450, loss: 0.00054535100935027
step: 460, loss: 5.0027283577946946e-05
step: 470, loss: 3.235296026105061e-05
step: 480, loss: 0.0006292458856478333
step: 490, loss: 0.0038481310475617647
step: 500, loss: 8.181606972357258e-05
step: 510, loss: 9.170754492515698e-05
step: 520, loss: 3.776877565542236e-05
step: 530, loss: 0.00010417146404506639
/usr/local/lib/python3.7/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=0.9287377736376339, f1=0.9281105990783408,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.488132203230634e-05
step: 10, loss: 0.00012692585005424917
step: 20, loss: 0.0009369656327180564
step: 30, loss: 8.613162935944274e-05
step: 40, loss: 0.0001599260576767847
step: 50, loss: 0.004954658914357424
step: 60, loss: 0.0003340544062666595
step: 70, loss: 0.00011038055527023971
step: 80, loss: 0.002533254213631153
step: 90, loss: 6.904316978761926e-05
step: 100, loss: 6.39909558231011e-05
step: 110, loss: 5.1263963541714475e-05
step: 120, loss: 4.766578058479354e-05
step: 130, loss: 4.076025288668461e-05
step: 140, loss: 9.591050184099004e-05
step: 150, loss: 9.502896864432842e-05
step: 160, loss: 4.4827276724390686e-05
step: 170, loss: 0.00029609212651848793
step: 180, loss: 4.45375480921939e-05
step: 190, loss: 0.00017309242684859782
step: 200, loss: 0.005945374257862568
step: 210, loss: 0.06909745931625366

```

```

step: 220, loss: 0.00019735825480893254
step: 230, loss: 0.015070823952555656
step: 240, loss: 4.938932761433534e-05
step: 250, loss: 4.630231705959886e-05
step: 260, loss: 2.913048410846386e-05
step: 270, loss: 4.207243910059333e-05
step: 280, loss: 6.648527050856501e-05
step: 290, loss: 0.00012593473365996033
step: 300, loss: 0.00047578534577041864
step: 310, loss: 3.50869850080926e-05
step: 320, loss: 0.0006554134888574481
step: 330, loss: 0.00018525983614381403
step: 340, loss: 0.00011880305100930855
step: 350, loss: 0.0019381908932700753
step: 360, loss: 9.413965744897723e-05
step: 370, loss: 0.0002890257746912539
step: 380, loss: 0.00474324868991971
step: 390, loss: 0.12543156743049622
step: 400, loss: 9.718396177049726e-05
step: 410, loss: 0.003172505646944046
step: 420, loss: 0.0018104633782058954
step: 430, loss: 4.798352892976254e-05
step: 440, loss: 0.001406183815561235
step: 450, loss: 0.003009216161444783
step: 460, loss: 3.8260561268543825e-05
step: 470, loss: 5.1451599574647844e-05
step: 480, loss: 5.098774272482842e-05
step: 490, loss: 0.002089925343170762
step: 500, loss: 0.0009100725292228162
step: 510, loss: 3.309838939458132e-05
step: 520, loss: 0.0002035644429270178
step: 530, loss: 0.000849552103318274
/usr/local/lib/python3.7/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=0.925627664613927, f1=0.9220657276995305,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.4496660848381e-05
step: 10, loss: 0.00020285529899410903
step: 20, loss: 8.613239333499223e-05
step: 30, loss: 0.0003022489254362881
step: 40, loss: 6.013663733028807e-05
step: 50, loss: 0.06969325244426727
step: 60, loss: 0.005524700973182917
step: 70, loss: 7.313566311495379e-05

```

step: 80, loss: 5.38286694791168e-05
step: 90, loss: 0.0001255209936061874
step: 100, loss: 0.000447086407802999
step: 110, loss: 0.0011851423187181354
step: 120, loss: 0.00010116400517290458
step: 130, loss: 5.7859138905769214e-05
step: 140, loss: 0.0006893788813613355
step: 150, loss: 3.5399523767409846e-05
step: 160, loss: 7.182167610153556e-05
step: 170, loss: 0.0002114870585501194
step: 180, loss: 0.00020899246737826616
step: 190, loss: 0.002458001719787717
step: 200, loss: 0.16616566479206085
step: 210, loss: 0.002611592411994934
step: 220, loss: 0.0052402731962502
step: 230, loss: 6.052759999874979e-05
step: 240, loss: 5.151806544745341e-05
step: 250, loss: 0.0016202987171709538
step: 260, loss: 0.0003083283663727343
step: 270, loss: 0.00010149353329325095
step: 280, loss: 7.91149286669679e-05
step: 290, loss: 0.00010242829011986032
step: 300, loss: 0.0008700749021954834
step: 310, loss: 0.0001873227156465873
step: 320, loss: 8.690552203916013e-05
step: 330, loss: 0.00034922861959785223
step: 340, loss: 0.00010747668420663103
step: 350, loss: 0.0005811388837173581
step: 360, loss: 0.00010185461724177003
step: 370, loss: 0.0010266718454658985
step: 380, loss: 0.001308142440393567
step: 390, loss: 0.00012888455239590257
step: 400, loss: 0.00016933314327616245
step: 410, loss: 0.0002705340157262981
step: 420, loss: 0.00036322977393865585
step: 430, loss: 8.748741674935445e-05
step: 440, loss: 0.0003254142648074776
step: 450, loss: 4.9427413614466786e-05
step: 460, loss: 0.020163284614682198
step: 470, loss: 0.00010275278327753767
step: 480, loss: 6.653767923125997e-05
step: 490, loss: 5.419378067017533e-05
step: 500, loss: 0.0002582525194156915
step: 510, loss: 0.000567534239962697
step: 520, loss: 7.685678428970277e-05
step: 530, loss: 0.0001023117802105844
/usr/local/lib/python3.7/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=0.9296693060083837, f1=0.9292929292929294,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001934905885718763
step: 10, loss: 2.981622310471721e-05
step: 20, loss: 4.2073894292116165e-05
step: 30, loss: 0.007871664129197598
step: 40, loss: 4.244661613483913e-05
step: 50, loss: 0.11473555117845535
step: 60, loss: 0.00014399905921891332
step: 70, loss: 0.0003922584292013198
step: 80, loss: 9.094107372220606e-05
step: 90, loss: 5.847337160957977e-05
step: 100, loss: 5.986632459098473e-05
step: 110, loss: 5.270614201435819e-05
step: 120, loss: 3.361921335454099e-05
step: 130, loss: 5.9903599321842194e-05
step: 140, loss: 0.00011356551840435714
step: 150, loss: 3.9545135223306715e-05
step: 160, loss: 3.7042507756268606e-05
step: 170, loss: 5.0007707613985986e-05
step: 180, loss: 0.0008311179699376225
step: 190, loss: 0.0006523139891214669
step: 200, loss: 0.00010870672849705443
step: 210, loss: 0.00016453812713734806
step: 220, loss: 7.653844659216702e-05
step: 230, loss: 2.5044313588296063e-05
step: 240, loss: 0.0005914937355555594
step: 250, loss: 6.418128032237291e-05
step: 260, loss: 0.0006053094984963536
step: 270, loss: 0.00018458378326613456
step: 280, loss: 0.0007565412088297307
step: 290, loss: 0.0008344292873516679
step: 300, loss: 0.0064218188636004925
step: 310, loss: 8.420886297244579e-05
step: 320, loss: 3.0318369681481272e-05
step: 330, loss: 0.00012755070929415524
step: 340, loss: 5.6181615946115926e-05
step: 350, loss: 0.001746694091707468
step: 360, loss: 0.0004753282410092652
step: 370, loss: 0.00017792581638786942
step: 380, loss: 4.338669532444328e-05
step: 390, loss: 5.9402718761703e-05
step: 400, loss: 3.9567865314893425e-05
step: 410, loss: 0.0012408335460349917
```

```

step: 420, loss: 0.000849634874612093
step: 430, loss: 6.700298399664462e-05
step: 440, loss: 0.0002594749676063657
step: 450, loss: 0.0001294472604058683
step: 460, loss: 5.258505552774295e-05
step: 470, loss: 3.304156780359335e-05
step: 480, loss: 6.115459109423682e-05
step: 490, loss: 9.270611190004274e-05
step: 500, loss: 3.6602305044652894e-05
step: 510, loss: 0.0001266083272639662
step: 520, loss: 5.306044113240205e-05
step: 530, loss: 3.818616096395999e-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=0.9285042333019755, f1=0.9233644859813084,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042584488983266056
step: 10, loss: 2.4083426978904754e-05
step: 20, loss: 5.134391176397912e-05
step: 30, loss: 0.0006645920220762491
step: 40, loss: 3.732127515831962e-05
step: 50, loss: 3.771621049963869e-05
step: 60, loss: 0.00011021124373655766
step: 70, loss: 4.280228677089326e-05
step: 80, loss: 4.324016845203005e-05
step: 90, loss: 0.0056884028017520905
step: 100, loss: 0.0010204642312601209
step: 110, loss: 2.4891229259083048e-05
step: 120, loss: 5.510899791261181e-05
step: 130, loss: 2.7709817004506476e-05
step: 140, loss: 4.002332570962608e-05
step: 150, loss: 0.00010962285159621388
step: 160, loss: 2.5368155547766946e-05
step: 170, loss: 3.6099903809372336e-05
step: 180, loss: 0.0002320111234439537
step: 190, loss: 2.8646414648392238e-05
step: 200, loss: 3.718951847986318e-05
step: 210, loss: 3.500353705021553e-05
step: 220, loss: 2.4209681214415468e-05
step: 230, loss: 3.127586387563497e-05
step: 240, loss: 5.1934559451183304e-05
step: 250, loss: 2.422447687422391e-05
step: 260, loss: 1.8316844943910837e-05
step: 270, loss: 2.362495615670923e-05

```

```
step: 280, loss: 0.00011123414151370525
step: 290, loss: 0.00014766253298148513
step: 300, loss: 3.2099487725645304e-05
step: 310, loss: 3.109021781710908e-05
step: 320, loss: 0.0006267015705816448
step: 330, loss: 5.631853855447844e-05
step: 340, loss: 3.274295522714965e-05
step: 350, loss: 3.172329525114037e-05
step: 360, loss: 3.3279735362157226e-05
step: 370, loss: 0.00016232846246566623
step: 380, loss: 7.738464773865417e-05
step: 390, loss: 0.00029292181716300547
step: 400, loss: 5.583492020377889e-05
step: 410, loss: 3.084369745920412e-05
step: 420, loss: 2.9726679713348858e-05
step: 430, loss: 2.3427815904142335e-05
step: 440, loss: 0.00025431503308936954
step: 450, loss: 6.079351078369655e-05
step: 460, loss: 7.436991290887818e-05
step: 470, loss: 1.7787981050787494e-05
step: 480, loss: 2.4783219487289898e-05
step: 490, loss: 3.6084682506043464e-05
step: 500, loss: 1.4763172657694668e-05
step: 510, loss: 0.003179767169058323
step: 520, loss: 0.0006601123022846878
step: 530, loss: 2.610215415188577e-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=0.9291784702549575, f1=0.9266012155212716,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009658378548920155
step: 10, loss: 2.3934020646265708e-05
step: 20, loss: 2.4191362172132358e-05
step: 30, loss: 2.4653256332385354e-05
step: 40, loss: 3.0781062378082424e-05
step: 50, loss: 6.790015322621912e-05
step: 60, loss: 0.0001021752177621238
step: 70, loss: 0.0005049858591519296
step: 80, loss: 3.7756235542474315e-05
step: 90, loss: 4.5204047637525946e-05
step: 100, loss: 5.769464769400656e-05
step: 110, loss: 3.446215850999579e-05
step: 120, loss: 8.079813414951786e-05
step: 130, loss: 0.0026856856420636177
```



```

step: 140, loss: 2.965270687127486e-05
step: 150, loss: 1.878619696071837e-05
step: 160, loss: 2.576257429609541e-05
step: 170, loss: 3.0490129574900493e-05
step: 180, loss: 0.0004011005803477019
step: 190, loss: 3.0072766094235703e-05
step: 200, loss: 8.288129902211949e-05
step: 210, loss: 2.6787807655637152e-05
step: 220, loss: 6.293628393905237e-05
step: 230, loss: 5.352908920031041e-05
step: 240, loss: 3.055707202292979e-05
step: 250, loss: 2.633654912642669e-05
step: 260, loss: 4.568460281006992e-05
step: 270, loss: 4.1051669541047886e-05
step: 280, loss: 1.9389741282793693e-05
step: 290, loss: 0.0025409541558474302
step: 300, loss: 3.365433803992346e-05
step: 310, loss: 0.0003108005039393902
step: 320, loss: 2.377051714574918e-05
step: 330, loss: 3.5530320019461215e-05
step: 340, loss: 0.0007014317088760436
step: 350, loss: 4.991194873582572e-05
step: 360, loss: 0.000187998783076182
step: 370, loss: 3.0046823667362332e-05
step: 380, loss: 4.944284592056647e-05
step: 390, loss: 0.0018902652664110065
step: 400, loss: 4.983653707313351e-05
step: 410, loss: 2.1442077922984026e-05
step: 420, loss: 2.6526147848926485e-05
step: 430, loss: 5.698877066606656e-05
step: 440, loss: 4.6882818423910066e-05
step: 450, loss: 3.606680184020661e-05
step: 460, loss: 2.6891244488069788e-05
step: 470, loss: 2.0660076188505627e-05
step: 480, loss: 2.1967416614643298e-05
step: 490, loss: 2.7573330953600816e-05
step: 500, loss: 2.4194772777264006e-05
step: 510, loss: 2.5681354600237682e-05
step: 520, loss: 1.7452710380894132e-05
step: 530, loss: 0.00038281892193481326
/usr/local/lib/python3.7/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=0.9284043051006083, f1=0.924860853432282,
best_f1=0.9305936073059361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 9.47099324548617e-05
step: 10, loss: 7.10505701135844e-05
step: 20, loss: 2.3442324163625017e-05
step: 30, loss: 5.634933040710166e-05
step: 40, loss: 0.021144209429621696
step: 50, loss: 4.015072408947162e-05
step: 60, loss: 1.4405509318748955e-05
step: 70, loss: 0.00012151000555604696
step: 80, loss: 2.296962520631496e-05
step: 90, loss: 3.818473487626761e-05
step: 100, loss: 7.1965747338254e-05
step: 110, loss: 3.880593430949375e-05
step: 120, loss: 0.00030694640008732677
step: 130, loss: 0.027915023267269135
step: 140, loss: 1.95796019397676e-05
step: 150, loss: 3.0594237614423037e-05
step: 160, loss: 2.2965365133131854e-05
step: 170, loss: 1.8737835489446297e-05
step: 180, loss: 1.4144776287139393e-05
step: 190, loss: 2.5729983462952077e-05
step: 200, loss: 3.219993595848791e-05
step: 210, loss: 3.470080264378339e-05
step: 220, loss: 2.5550010832375847e-05
step: 230, loss: 0.0008796838228590786
step: 240, loss: 2.2034600988263264e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 250, loss: 2.835589839378372e-05
step: 260, loss: 2.874708479794208e-05
step: 270, loss: 1.8868286133510992e-05
step: 280, loss: 2.6228670321870595e-05
step: 290, loss: 1.5970084859873168e-05
step: 300, loss: 2.2280184566625394e-05
step: 310, loss: 0.00011468303273431957
step: 320, loss: 0.0007549577858299017
step: 330, loss: 5.2822986617684364e-05
step: 340, loss: 1.6391039025620557e-05
step: 350, loss: 1.4047903277969453e-05
step: 360, loss: 0.004086992237716913
step: 370, loss: 3.576728340703994e-05
step: 380, loss: 2.2518981495522894e-05
step: 390, loss: 3.58946890628431e-05
step: 400, loss: 2.201603820140008e-05
step: 410, loss: 2.523373404983431e-05
step: 420, loss: 2.6075418645632453e-05
step: 430, loss: 3.51647031493485e-05
step: 440, loss: 0.0004808813682757318
step: 450, loss: 1.7258920706808567e-05
step: 460, loss: 1.9940986021538265e-05

```

step: 470, loss: 0.009627853520214558
step: 480, loss: 1.4822748198639601e-05
step: 490, loss: 1.585461177455727e-05
step: 500, loss: 1.8261061995872296e-05
step: 510, loss: 3.0144909032969736e-05
step: 520, loss: 1.9620665625552647e-05
step: 530, loss: 1.6871606931090355e-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=0.9285714285714286, f1=0.9252206223873664,
best_f1=0.9305936073059361
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, 254.35it/s]
load_f1 = 0.9332105020727776
real_f1 = 0.9313680331644404
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:17, 258.71it/s]

4 DITTO

4.1 DITTO STRUCTURED

4.1.1 Beer - Running the matcher

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

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

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

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

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5312608480453491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5116279069767441, f1=0.37499999999999994,
best_f1=0.37499999999999994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4874518811702728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5384615384615384, f1=0.43750000000000006,
best_f1=0.43750000000000006
/usr/local/lib/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 16384.0
step: 0, loss: 0.6368667483329773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6470588235294117, f1=0.4285714285714286,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3194195032119751
/usr/local/lib/python3.7/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=0.6206896551724138, f1=0.5, best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24713538587093353
/usr/local/lib/python3.7/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=0.5925925925925927, f1=0.6086956521739131,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2252349853515625
/usr/local/lib/python3.7/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=0.4444444444444445, f1=0.5, best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25485286116600037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.7741935483870968, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.046492867171764374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8, f1=0.6666666666666666, best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0090756481513381
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0735405907034874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8571428571428571, 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: 0.010762370191514492
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.7142857142857143,
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: 0.008753449656069279
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.7142857142857143,
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: 0.007149437442421913
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.689655172413793,
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: 0.014762338250875473
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.689655172413793,
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: 0.02753913588821888
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.689655172413793,
best_f1=0.6428571428571429
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled          : True

```



```

opt_level          : 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, 138140.31it/s]
/usr/local/lib/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 = 0.7692307692307692
real_f1 = 0.7142857142857143
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 252.99it/s]

```

4.1.2 DBLP-ACM - Running the matcher

```

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

```

```

--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

An input tensor was not cuda.

```

    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6119521260261536
step: 10, loss: 0.6320194005966187
step: 20, loss: 0.3966062068939209
step: 30, loss: 0.18395905196666718
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.13047753274440765
step: 50, loss: 0.11584755033254623
step: 60, loss: 0.13069503009319305
step: 70, loss: 0.14997518062591553
step: 80, loss: 0.05001157522201538
step: 90, loss: 0.16340090334415436
step: 100, loss: 0.03863667696714401
step: 110, loss: 0.10968505591154099
step: 120, loss: 0.032624583691358566
step: 130, loss: 0.004386653192341328
step: 140, loss: 0.0031239406671375036
step: 150, loss: 0.06306209415197372
step: 160, loss: 0.008205048739910126
step: 170, loss: 0.008614420890808105
step: 180, loss: 0.04674219340085983
step: 190, loss: 0.02737077698111534
step: 200, loss: 0.032720062881708145
step: 210, loss: 0.03903074562549591
step: 220, loss: 0.012806260958313942
step: 230, loss: 0.020425893366336823
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776286353467561, f1=0.9819413092550789,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.009570679627358913
step: 10, loss: 0.0021702637895941734
step: 20, loss: 0.13749030232429504
step: 30, loss: 0.20359547436237335
step: 40, loss: 0.041633788496255875
step: 50, loss: 0.006354347337037325
step: 60, loss: 0.0031593809835612774
step: 70, loss: 0.12057261914014816
step: 80, loss: 0.003240276360884309
step: 90, loss: 0.007934567518532276
step: 100, loss: 0.02390563115477562
step: 110, loss: 0.11772334575653076
step: 120, loss: 0.03494066372513771
step: 130, loss: 0.002775535685941577
step: 140, loss: 0.0008664231863804162
step: 150, loss: 0.14031986892223358
step: 160, loss: 0.0183248370885849
step: 170, loss: 0.0034952715504914522
step: 180, loss: 0.00426105409860611
step: 190, loss: 0.012890639714896679
step: 200, loss: 0.0016118843341246247
step: 210, loss: 0.0009349846513941884
step: 220, loss: 0.04310249537229538
step: 230, loss: 0.004719787277281284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9783845278725825, f1=0.9690011481056257,
best_f1=0.9690011481056257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009824671782553196
step: 10, loss: 0.04160129651427269
step: 20, loss: 0.0055526550859212875
step: 30, loss: 0.02440972439944744
step: 40, loss: 0.1441362500190735
step: 50, loss: 0.021404653787612915
step: 60, loss: 0.009233182296156883
step: 70, loss: 0.05258650705218315
step: 80, loss: 0.0024872669018805027
step: 90, loss: 0.24224799871444702
step: 100, loss: 0.0005976793472655118
step: 110, loss: 0.003051646752282977

```

```

step: 120, loss: 0.017747540026903152
step: 130, loss: 0.0038635896053165197
step: 140, loss: 0.0769595056772232
step: 150, loss: 0.010087411850690842
step: 160, loss: 0.029909078031778336
step: 170, loss: 0.006087037734687328
step: 180, loss: 0.044811517000198364
step: 190, loss: 0.001473727053962648
step: 200, loss: 0.009372128173708916
step: 210, loss: 0.02651895396411419
step: 220, loss: 0.0002769209386315197
step: 230, loss: 0.0006882900488562882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831649831649831, f1=0.9772727272727272,
best_f1=0.9772727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017649215878918767
step: 10, loss: 0.0004579265951178968
step: 20, loss: 0.002430754480883479
step: 30, loss: 0.0017405081307515502
step: 40, loss: 0.00442076101899147
step: 50, loss: 0.0013839834136888385
step: 60, loss: 0.003266040002927184
step: 70, loss: 0.030698223039507866
step: 80, loss: 0.003081359900534153
step: 90, loss: 0.00022231711773201823
step: 100, loss: 0.0003345714067108929
step: 110, loss: 0.0002690369146876037
step: 120, loss: 0.03802518919110298
step: 130, loss: 0.006016362924128771
step: 140, loss: 0.0017398486379534006
step: 150, loss: 0.0700223445892334
step: 160, loss: 0.002442996483296156
step: 170, loss: 0.02770053967833519
step: 180, loss: 0.0001924084936035797
step: 190, loss: 0.004833288956433535
step: 200, loss: 0.0001731912634568289
step: 210, loss: 0.23761853575706482
step: 220, loss: 0.0002352128503844142
step: 230, loss: 0.09027975797653198

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842696629213483, 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: 0.0023607334587723017
step: 10, loss: 0.000852513883728534
step: 20, loss: 0.011723445728421211
step: 30, loss: 0.000194321182789281
step: 40, loss: 0.0005154898972250521
step: 50, loss: 0.0011893854243680835
step: 60, loss: 0.0007133638137020171
step: 70, loss: 0.00020972391939722002
step: 80, loss: 0.00039949099300429225
step: 90, loss: 0.00021362316329032183
step: 100, loss: 0.0010471612913534045
step: 110, loss: 0.00016984163084998727
step: 120, loss: 0.00011168736091349274
step: 130, loss: 0.0004266660544089973
step: 140, loss: 0.0007852716953493655
step: 150, loss: 0.0007702477741986513
step: 160, loss: 0.00035007225233130157
step: 170, loss: 0.002594652818515897
step: 180, loss: 0.0006106546497903764
step: 190, loss: 0.0006160433986224234
step: 200, loss: 0.002288282150402665
step: 210, loss: 0.043588414788246155
step: 220, loss: 0.0020706134382635355
step: 230, loss: 0.005520595237612724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9886877828054299, f1=0.9783845278725825,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0022845056373625994
step: 10, loss: 9.073293040273711e-05
step: 20, loss: 0.0014610482612624764
step: 30, loss: 0.0010932632721960545
step: 40, loss: 0.0013287428300827742
step: 50, loss: 0.004272374790161848
step: 60, loss: 0.00010571171878837049
step: 70, loss: 0.0007537496858276427
step: 80, loss: 0.0017259283922612667
step: 90, loss: 0.0007989482255652547
step: 100, loss: 0.009035227820277214
step: 110, loss: 0.008809835650026798
step: 120, loss: 0.014923079870641232
step: 130, loss: 0.003867285791784525
step: 140, loss: 0.00019508867990225554
step: 150, loss: 0.0002501648268662393
step: 160, loss: 0.0011931016342714429
step: 170, loss: 0.003797756275162101
step: 180, loss: 0.0279773510992527
step: 190, loss: 0.006227744277566671
step: 200, loss: 0.0010416826698929071
step: 210, loss: 0.001936302985996008
step: 220, loss: 8.07526012067683e-05
step: 230, loss: 0.17755484580993652
/usr/local/lib/python3.7/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=0.9863945578231292, f1=0.9748283752860413,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001831941888667643
step: 10, loss: 0.0002754195884335786
step: 20, loss: 0.0018335244385525584
step: 30, loss: 0.011233076453208923
step: 40, loss: 0.004030771087855101
step: 50, loss: 0.00032303648185916245
step: 60, loss: 0.034265726804733276
step: 70, loss: 0.030552668496966362
step: 80, loss: 0.00010625545110087842
step: 90, loss: 0.013238289393484592
step: 100, loss: 0.00017693791596684605
step: 110, loss: 0.006752906367182732
step: 120, loss: 0.0001496905751992017
step: 130, loss: 0.00029072928009554744
step: 140, loss: 0.0001647027675062418
step: 150, loss: 0.0002551142533775419

```

```

step: 160, loss: 0.015057210810482502
step: 170, loss: 0.0023227743804454803
step: 180, loss: 0.0002328577684238553
step: 190, loss: 0.00021302406094036996
step: 200, loss: 0.11806193739175797
step: 210, loss: 9.28108929656446e-05
step: 220, loss: 0.04271569475531578
step: 230, loss: 0.0009719508816488087
/usr/local/lib/python3.7/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=0.9819004524886877, f1=0.9714285714285715,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010602332622511312
step: 10, loss: 0.0030575005803257227
step: 20, loss: 0.0012956878636032343
step: 30, loss: 0.0002969769120682031
step: 40, loss: 0.0014297675807029009
step: 50, loss: 0.0020764300134032965
step: 60, loss: 0.00015977027942426503
step: 70, loss: 0.0016202026745304465
step: 80, loss: 0.00676490506157279
step: 90, loss: 0.0013715948443859816
step: 100, loss: 7.764493784634396e-05
step: 110, loss: 0.004911927971988916
step: 120, loss: 0.00022161621018312871
step: 130, loss: 0.002368551678955555
step: 140, loss: 8.040050306590274e-05
step: 150, loss: 8.193840767489746e-05
step: 160, loss: 0.00019525413517840207
step: 170, loss: 0.00012564957432914525
step: 180, loss: 0.00031568598933517933
step: 190, loss: 0.002331459429115057
step: 200, loss: 8.646618516650051e-05
step: 210, loss: 0.00021437305258587003
step: 220, loss: 0.0001451491261832416
step: 230, loss: 6.297662912402302e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9841986455981941, f1=0.9783845278725825,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.03511499986052513
step: 10, loss: 0.00014767654647585005
step: 20, loss: 0.003114062827080488
step: 30, loss: 5.062039053882472e-05
step: 40, loss: 0.011223874054849148
step: 50, loss: 3.5333490814082325e-05
step: 60, loss: 0.0004112793831154704
step: 70, loss: 7.912566070444882e-05
step: 80, loss: 4.267531767254695e-05
step: 90, loss: 5.8858568081632257e-05
step: 100, loss: 4.555869963951409e-05
step: 110, loss: 3.177601320203394e-05
step: 120, loss: 4.131568857701495e-05
step: 130, loss: 7.999210356501862e-05
step: 140, loss: 4.339758743299171e-05
step: 150, loss: 7.618516974616796e-05
step: 160, loss: 0.00011734799772966653
step: 170, loss: 0.0009921913733705878
step: 180, loss: 0.0012631412828341126
step: 190, loss: 8.19471460999921e-05
step: 200, loss: 0.000665616593323648
step: 210, loss: 0.0061089531518518925
step: 220, loss: 0.0028803213499486446
step: 230, loss: 6.232860323507339e-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=0.983050847457627, f1=0.9749430523917996,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.06566975673195e-05
step: 10, loss: 3.0494420570903458e-05
step: 20, loss: 4.054389864904806e-05
step: 30, loss: 0.0011058928212150931
step: 40, loss: 6.3886254793033e-05
step: 50, loss: 0.00022239997633732855
step: 60, loss: 0.0025838392321020365
step: 70, loss: 0.0012767630396410823
step: 80, loss: 0.00020674726692959666
step: 90, loss: 0.0003407710464671254
step: 100, loss: 5.721994966734201e-05
step: 110, loss: 0.00011954367073485628
step: 120, loss: 0.0005900111282244325
step: 130, loss: 0.00015871501818764955
step: 140, loss: 0.00011341203207848594
step: 150, loss: 0.006286763586103916

```

```

step: 160, loss: 4.415634975885041e-05
step: 170, loss: 0.00021117986761964858
step: 180, loss: 0.0008018256048671901
step: 190, loss: 0.00016408869123551995
step: 200, loss: 4.430478657013737e-05
step: 210, loss: 3.924715565517545e-05
step: 220, loss: 7.530902075814083e-05
step: 230, loss: 0.00014659855514764786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9831271091113611, f1=0.9807037457434733,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.3113468311494216e-05
step: 10, loss: 9.831094939727336e-05
step: 20, loss: 0.0006329728057608008
step: 30, loss: 0.06321541965007782
step: 40, loss: 0.005129561759531498
step: 50, loss: 0.00015125393110793084
step: 60, loss: 0.0002299566986039281
step: 70, loss: 0.00026566695305518806
step: 80, loss: 0.00015012131188996136
step: 90, loss: 9.152062557404861e-05
step: 100, loss: 0.0001627151359571144
step: 110, loss: 0.00997261330485344
step: 120, loss: 7.990459562279284e-05
step: 130, loss: 0.006434336304664612
step: 140, loss: 0.0006575208972208202
step: 150, loss: 0.017074210569262505
step: 160, loss: 0.00011697770969476551
step: 170, loss: 0.02353673242032528
step: 180, loss: 0.00037564741796813905
step: 190, loss: 0.00011912012996617705
step: 200, loss: 0.0023166879545897245
step: 210, loss: 4.1457489714957774e-05
step: 220, loss: 4.261529466020875e-05
step: 230, loss: 0.0014469185844063759
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9748858447488584,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 5.295917435432784e-05
step: 10, loss: 0.00010867956734728068
step: 20, loss: 4.7044562961673364e-05
step: 30, loss: 6.25488391960971e-05
step: 40, loss: 9.344983845949173e-05
step: 50, loss: 0.0002633772965054959
step: 60, loss: 0.000152588399942033
step: 70, loss: 4.570201053866185e-05
step: 80, loss: 0.0002382782695349306
step: 90, loss: 3.082199691561982e-05
step: 100, loss: 3.143228605040349e-05
step: 110, loss: 0.0023129417095333338
step: 120, loss: 0.00024710275465622544
step: 130, loss: 0.0009764139540493488
step: 140, loss: 0.00042782517266459763
step: 150, loss: 8.712820999789983e-05
step: 160, loss: 4.814334533875808e-05
step: 170, loss: 0.00022618795628659427
step: 180, loss: 0.0043862853199243546
step: 190, loss: 0.00011693139822455123
step: 200, loss: 0.00011406250996515155
step: 210, loss: 0.00039137370185926557
step: 220, loss: 0.0004330161027610302
step: 230, loss: 0.004381218459457159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9842342342342343, f1=0.9807037457434733,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008928145980462432
step: 10, loss: 0.0010743738384917378
step: 20, loss: 0.0002228631783509627
step: 30, loss: 0.0001497122721048072
step: 40, loss: 0.00017051008762791753
step: 50, loss: 4.632107084034942e-05
step: 60, loss: 3.83196456823498e-05
step: 70, loss: 2.669084642548114e-05
step: 80, loss: 7.762091991025954e-05
step: 90, loss: 0.0035624862648546696
step: 100, loss: 1.8901877410826273e-05
step: 110, loss: 0.0007700494606979191
step: 120, loss: 0.03729894384741783
step: 130, loss: 4.238696419633925e-05
step: 140, loss: 0.0004997371579520404
step: 150, loss: 2.8233318516868167e-05

```

```

step: 160, loss: 4.72850697406102e-05
step: 170, loss: 5.303515717969276e-05
step: 180, loss: 5.8702782553154975e-05
step: 190, loss: 2.396781746938359e-05
step: 200, loss: 0.000152592605445534
step: 210, loss: 1.7180800568894483e-05
step: 220, loss: 0.00035892106825485826
step: 230, loss: 3.4081953344866633e-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=0.9864253393665158, f1=0.9738933030646991,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.4024160413537174e-05
step: 10, loss: 0.00016864936333149672
step: 20, loss: 8.243475895142183e-05
step: 30, loss: 0.00011313323921058327
step: 40, loss: 0.05280320718884468
step: 50, loss: 3.074764026678167e-05
step: 60, loss: 0.0005669296369887888
step: 70, loss: 0.0009028776548802853
step: 80, loss: 4.333343531470746e-05
step: 90, loss: 0.0007009063847362995
step: 100, loss: 2.8955431844224222e-05
step: 110, loss: 0.00038334738928824663
step: 120, loss: 5.1957526011392474e-05
step: 130, loss: 7.348440703935921e-05
step: 140, loss: 8.972292562248185e-05
step: 150, loss: 0.00023403919476550072
step: 160, loss: 0.00010173532064072788
step: 170, loss: 3.77277574443724e-05
step: 180, loss: 4.773679393110797e-05
step: 190, loss: 3.215908873244189e-05
step: 200, loss: 2.526795469748322e-05
step: 210, loss: 7.120895315892994e-05
step: 220, loss: 0.0002629757800605148
step: 230, loss: 0.00016967864939942956
/usr/local/lib/python3.7/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=0.9864559819413092, f1=0.9784335981838819,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 2.53015969065018e-05
step: 10, loss: 1.855163100117352e-05
step: 20, loss: 7.92898572399281e-05
step: 30, loss: 6.628797564189881e-05
step: 40, loss: 4.325263944338076e-05
step: 50, loss: 0.00246491190046072
step: 60, loss: 2.677297015907243e-05
step: 70, loss: 2.7059815693064593e-05
step: 80, loss: 9.15186174097471e-05
step: 90, loss: 7.93967628851533e-05
step: 100, loss: 5.007999061490409e-05
step: 110, loss: 1.8812459529726766e-05
step: 120, loss: 0.00025368938804604113
step: 130, loss: 0.0008557206019759178
step: 140, loss: 2.9786237064399756e-05
step: 150, loss: 0.0001924981625052169
step: 160, loss: 2.243678864033427e-05
step: 170, loss: 2.4228251277236268e-05
step: 180, loss: 0.000944113009609282
step: 190, loss: 0.0003310814208816737
step: 200, loss: 0.0008527630707249045
step: 210, loss: 4.116435957257636e-05
step: 220, loss: 2.322302316315472e-05
step: 230, loss: 2.2351361621986143e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9864559819413092, f1=0.9784335981838819,
best_f1=0.9783845278725825
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 198.30it/s]
load_f1 = 0.9853107344632768
real_f1 = 0.9841986455981941
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 243.74it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

# Step 3: run the matcher

```

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

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: 0.6027225852012634
```

```
step: 10, loss: 0.5248605608940125
```

step: 20, loss: 0.5609254837036133
step: 30, loss: 0.27366483211517334
step: 40, loss: 0.16865845024585724
step: 50, loss: 0.2864857614040375
step: 60, loss: 0.07524896413087845
step: 70, loss: 0.23246632516384125
step: 80, loss: 0.03317463770508766
step: 90, loss: 0.45145076513290405
step: 100, loss: 0.1075524240732193
step: 110, loss: 0.11142952740192413
step: 120, loss: 0.1764335185289383
step: 130, loss: 0.06150165945291519
step: 140, loss: 0.11427386850118637
step: 150, loss: 0.06821471452713013
step: 160, loss: 0.04282798618078232
step: 170, loss: 0.29168736934661865
step: 180, loss: 0.14896640181541443
step: 190, loss: 0.008880237117409706
step: 200, loss: 0.08154404908418655
step: 210, loss: 0.06436330825090408
step: 220, loss: 0.24515847861766815
step: 230, loss: 0.21050721406936646
step: 240, loss: 0.06042942404747009
step: 250, loss: 0.06451636552810669
step: 260, loss: 0.12330003082752228
step: 270, loss: 0.011620786041021347
step: 280, loss: 0.06727845221757889
step: 290, loss: 0.024133747443556786
step: 300, loss: 0.05002298206090927
step: 310, loss: 0.21043454110622406
step: 320, loss: 0.13824914395809174
step: 330, loss: 0.0765714943408966
step: 340, loss: 0.16822023689746857
step: 350, loss: 0.06767553836107254
step: 360, loss: 0.14780515432357788
step: 370, loss: 0.1641445755958557
step: 380, loss: 0.031922318041324615
step: 390, loss: 0.15980055928230286
step: 400, loss: 0.3274146318435669
step: 410, loss: 0.07874012738466263
step: 420, loss: 0.05929026007652283
step: 430, loss: 0.1738949716091156
step: 440, loss: 0.04280742257833481
step: 450, loss: 0.03836021572351456
step: 460, loss: 0.01095914002507925
step: 470, loss: 0.1655518263578415
step: 480, loss: 0.0717751532793045
step: 490, loss: 0.08318477123975754


```

step: 500, loss: 0.053750328719615936
step: 510, loss: 0.12070059031248093
step: 520, loss: 0.0540330596268177
step: 530, loss: 0.004449714440852404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9221957040572791, f1=0.9113323850165956,
best_f1=0.9113323850165956
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18077009916305542
step: 10, loss: 0.0542944110929966
step: 20, loss: 0.035771846771240234
step: 30, loss: 0.050931841135025024
step: 40, loss: 0.03574242442846298
step: 50, loss: 0.11851680278778076
step: 60, loss: 0.016822574660182
step: 70, loss: 0.06978228688240051
step: 80, loss: 0.059401970356702805
step: 90, loss: 0.01948206126689911
step: 100, loss: 0.030022161081433296
step: 110, loss: 0.011295998468995094
step: 120, loss: 0.050667110830545425
step: 130, loss: 0.15113531053066254
step: 140, loss: 0.015199731104075909
step: 150, loss: 0.11218850314617157
step: 160, loss: 0.052897270768880844
step: 170, loss: 0.028358256444334984
step: 180, loss: 0.03632698580622673
step: 190, loss: 0.1553831398487091
step: 200, loss: 0.030713628977537155
step: 210, loss: 0.06392324715852737
step: 220, loss: 0.10743097960948944
step: 230, loss: 0.006921133492141962
step: 240, loss: 0.07112351059913635
step: 250, loss: 0.010219361633062363
step: 260, loss: 0.004413976799696684
step: 270, loss: 0.23708076775074005
step: 280, loss: 0.03884409740567207
step: 290, loss: 0.12660838663578033
step: 300, loss: 0.10062378644943237
step: 310, loss: 0.025978775694966316

```

```

step: 320, loss: 0.20858076214790344
step: 330, loss: 0.06800892949104309
step: 340, loss: 0.025841930881142616
step: 350, loss: 0.004137239884585142
step: 360, loss: 0.13714684545993805
step: 370, loss: 0.25077980756759644
step: 380, loss: 0.12159524112939835
step: 390, loss: 0.07194545120000839
step: 400, loss: 0.10242835432291031
step: 410, loss: 0.0892971009016037
step: 420, loss: 0.03838295862078667
step: 430, loss: 0.027993349358439445
step: 440, loss: 0.12038757652044296
step: 450, loss: 0.021386554464697838
step: 460, loss: 0.020359262824058533
step: 470, loss: 0.14895635843276978
step: 480, loss: 0.24773184955120087
step: 490, loss: 0.027301687747240067
step: 500, loss: 0.13266490399837494
step: 510, loss: 0.030127231031656265
step: 520, loss: 0.0756615474820137
step: 530, loss: 0.09631529450416565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9263746505125815, f1=0.9212855146716348,
best_f1=0.9212855146716348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08869315683841705
step: 10, loss: 0.010489525273442268
step: 20, loss: 0.02486037090420723
step: 30, loss: 0.08209104835987091
step: 40, loss: 0.005318774376064539
step: 50, loss: 0.10326119512319565
step: 60, loss: 0.04682067036628723
step: 70, loss: 0.010908488184213638
step: 80, loss: 0.045975152403116226
step: 90, loss: 0.008009064942598343
step: 100, loss: 0.004330598749220371
step: 110, loss: 0.010833692736923695
step: 120, loss: 0.0022055141162127256
step: 130, loss: 0.004388727247714996

```

```

step: 140, loss: 0.07261486351490021
step: 150, loss: 0.11522722244262695
step: 160, loss: 0.08908841013908386
step: 170, loss: 0.04270469769835472
step: 180, loss: 0.0319291315972805
step: 190, loss: 0.005155839957296848
step: 200, loss: 0.20084582269191742
step: 210, loss: 0.10066509991884232
step: 220, loss: 0.18799343705177307
step: 230, loss: 0.08225192874670029
step: 240, loss: 0.09955044090747833
step: 250, loss: 0.13539212942123413
step: 260, loss: 0.04476417228579521
step: 270, loss: 0.012943112291395664
step: 280, loss: 0.10126278549432755
step: 290, loss: 0.011521512642502785
step: 300, loss: 0.09737710654735565
step: 310, loss: 0.030750881880521774
step: 320, loss: 0.009759066626429558
step: 330, loss: 0.001834260649047792
step: 340, loss: 0.03873584792017937
step: 350, loss: 0.02043621614575386
step: 360, loss: 0.04109922796487808
step: 370, loss: 0.011703576892614365
step: 380, loss: 0.019817359745502472
step: 390, loss: 0.10630851238965988
step: 400, loss: 0.020518964156508446
step: 410, loss: 0.022242268547415733
step: 420, loss: 0.027269816026091576
step: 430, loss: 0.01757069304585457
step: 440, loss: 0.020258665084838867
step: 450, loss: 0.07971029728651047
step: 460, loss: 0.0502597838640213
step: 470, loss: 0.02216247096657753
step: 480, loss: 0.00661153020337224
step: 490, loss: 0.009779248386621475
step: 500, loss: 0.03477201238274574
step: 510, loss: 0.013776170089840889
step: 520, loss: 0.08079337328672409
step: 530, loss: 0.17300571501255035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9343339587242026, f1=0.9196597353497165,

```

```

best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01638600043952465
step: 10, loss: 0.025976553559303284
step: 20, loss: 0.0031114164739847183
step: 30, loss: 0.08196032792329788
step: 40, loss: 0.03824799507856369
step: 50, loss: 0.03999857231974602
step: 60, loss: 0.0023479368537664413
step: 70, loss: 0.003339284798130393
step: 80, loss: 0.021459408104419708
step: 90, loss: 0.1472693681716919
step: 100, loss: 0.007730613462626934
step: 110, loss: 0.1122080385684967
step: 120, loss: 0.018225960433483124
step: 130, loss: 0.022847622632980347
step: 140, loss: 0.006640251725912094
step: 150, loss: 0.011865685693919659
step: 160, loss: 0.0014445021515712142
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.008324642665684223
step: 180, loss: 0.026958690956234932
step: 190, loss: 0.007011327426880598
step: 200, loss: 0.0030369232408702374
step: 210, loss: 0.06537520885467529
step: 220, loss: 0.0024432369973510504
step: 230, loss: 0.012608624994754791
step: 240, loss: 0.00255486648529768
step: 250, loss: 0.13483616709709167
step: 260, loss: 0.0021701445803046227
step: 270, loss: 0.014279897324740887
step: 280, loss: 0.11072778701782227
step: 290, loss: 0.010756256058812141
step: 300, loss: 0.004549774341285229
step: 310, loss: 0.020180055871605873
step: 320, loss: 0.0047941249795258045
step: 330, loss: 0.009764477610588074
step: 340, loss: 0.2757898271083832
step: 350, loss: 0.0173221156001091
step: 360, loss: 0.16653884947299957
step: 370, loss: 0.031898725777864456
step: 380, loss: 0.002829926088452339
step: 390, loss: 0.06976434588432312
step: 400, loss: 0.0006932303658686578
step: 410, loss: 0.02760433591902256
step: 420, loss: 0.011051656678318977

```

```
step: 430, loss: 0.11764853447675705
step: 440, loss: 0.003594162641093135
step: 450, loss: 0.002771785482764244
step: 460, loss: 0.014180314727127552
step: 470, loss: 0.09907360374927521
step: 480, loss: 0.038224831223487854
step: 490, loss: 0.049055926501750946
step: 500, loss: 0.0045284610241651535
step: 510, loss: 0.0902346596121788
step: 520, loss: 0.23936715722084045
step: 530, loss: 0.036247774958610535
/usr/local/lib/python3.7/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=0.9232209737827715, f1=0.9255966307908282,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09508976340293884
step: 10, loss: 0.049347635358572006
step: 20, loss: 0.0020511087495833635
step: 30, loss: 0.0011965903686359525
step: 40, loss: 0.008802590891718864
step: 50, loss: 0.01768823340535164
step: 60, loss: 0.013535279780626297
step: 70, loss: 0.015293724834918976
step: 80, loss: 0.0003011561930179596
step: 90, loss: 0.010182740166783333
step: 100, loss: 0.03178233280777931
step: 110, loss: 0.08349765837192535
step: 120, loss: 0.0009176692110486329
step: 130, loss: 0.0004608291492331773
step: 140, loss: 0.0007364349439740181
step: 150, loss: 0.07571054250001907
step: 160, loss: 0.013278876431286335
step: 170, loss: 0.0755438581109047
step: 180, loss: 0.0015632956055924296
step: 190, loss: 0.0025540473870933056
step: 200, loss: 0.006317264400422573
step: 210, loss: 0.02109283022582531
step: 220, loss: 0.0016305005410686135
step: 230, loss: 0.05846444517374039
step: 240, loss: 0.030273541808128357
step: 250, loss: 0.007171341218054295
step: 260, loss: 0.010281945578753948
step: 270, loss: 0.0007076302426867187
step: 280, loss: 0.0053613027557730675
```

```

step: 290, loss: 0.1532503366470337
step: 300, loss: 0.006764579098671675
step: 310, loss: 0.015771469101309776
step: 320, loss: 0.007730277720838785
step: 330, loss: 0.0198633074760437
step: 340, loss: 0.014355208724737167
step: 350, loss: 0.0011675313580781221
step: 360, loss: 0.009327557869255543
step: 370, loss: 0.014738506637513638
step: 380, loss: 0.0013913502916693687
step: 390, loss: 0.0018760261591523886
step: 400, loss: 0.016377942636609077
step: 410, loss: 0.000984568614512682
step: 420, loss: 0.0014684437774121761
step: 430, loss: 0.08622609078884125
step: 440, loss: 0.025585250928997993
step: 450, loss: 0.0011002212995663285
step: 460, loss: 0.0008115798118524253
step: 470, loss: 0.0016818612348288298
step: 480, loss: 0.014119390398263931
step: 490, loss: 0.0009224622626788914
step: 500, loss: 0.06830549240112305
step: 510, loss: 0.42053741216659546
step: 520, loss: 0.04984995722770691
step: 530, loss: 0.0020773366559296846
/usr/local/lib/python3.7/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=0.9327014218009478, f1=0.9216896060749881,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008883113041520119
step: 10, loss: 0.0014933800557628274
step: 20, loss: 0.1199595257639885
step: 30, loss: 0.007510119583457708
step: 40, loss: 0.0007774441037327051
step: 50, loss: 0.0019807820208370686
step: 60, loss: 0.029301779344677925
step: 70, loss: 0.0007677313406020403
step: 80, loss: 0.011396870948374271
step: 90, loss: 0.0004648103204090148
step: 100, loss: 0.053590625524520874
step: 110, loss: 0.0045950328931212425
step: 120, loss: 0.0004950874717906117
step: 130, loss: 0.0004143494588788599
step: 140, loss: 0.04072875902056694

```

```

step: 150, loss: 0.004651033319532871
step: 160, loss: 0.0010971815790981054
step: 170, loss: 0.004250432830303907
step: 180, loss: 0.008119088597595692
step: 190, loss: 0.01200677640736103
step: 200, loss: 0.002810741076245904
step: 210, loss: 0.03843621909618378
step: 220, loss: 0.001625599805265665
step: 230, loss: 0.004654188174754381
step: 240, loss: 0.03287605196237564
step: 250, loss: 0.1406048685312271
step: 260, loss: 0.000456181209301576
step: 270, loss: 0.06995920836925507
step: 280, loss: 0.04051553085446358
step: 290, loss: 0.0005730538396164775
step: 300, loss: 0.003917717840522528
step: 310, loss: 0.0028159581124782562
step: 320, loss: 0.013918334618210793
step: 330, loss: 0.001207931898534298
step: 340, loss: 0.04641515761613846
step: 350, loss: 0.012558255344629288
step: 360, loss: 0.008676254190504551
step: 370, loss: 0.0021699857898056507
step: 380, loss: 0.0006655913894064724
step: 390, loss: 0.003809316549450159
step: 400, loss: 0.1776350736618042
step: 410, loss: 0.0006591595592908561
step: 420, loss: 0.016191717237234116
step: 430, loss: 0.005095874425023794
step: 440, loss: 0.004253785125911236
step: 450, loss: 0.02200189419090748
step: 460, loss: 0.007943516597151756
step: 470, loss: 0.009811618365347385
step: 480, loss: 0.03250354155898094
step: 490, loss: 0.004961377941071987
step: 500, loss: 0.000703439291100949
step: 510, loss: 0.12309662997722626
step: 520, loss: 0.016526347026228905
step: 530, loss: 0.011976123787462711
/usr/local/lib/python3.7/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=0.9271402550091076, f1=0.9162471395881007,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002421953249722719

```

step: 10, loss: 0.015606606379151344
step: 20, loss: 0.0006528085796162486
step: 30, loss: 0.0005889585008844733
step: 40, loss: 0.024917133152484894
step: 50, loss: 0.005814436823129654
step: 60, loss: 0.00264804158359766
step: 70, loss: 0.0032213127706199884
step: 80, loss: 0.0010414286516606808
step: 90, loss: 0.15661419928073883
step: 100, loss: 0.0006262148963287473
step: 110, loss: 0.0024621873162686825
step: 120, loss: 0.0031457922887057066
step: 130, loss: 0.003272402798756957
step: 140, loss: 0.0007555532502010465
step: 150, loss: 0.016651567071676254
step: 160, loss: 0.00045968921040184796
step: 170, loss: 0.0025318234693259
step: 180, loss: 0.0005637973081320524
step: 190, loss: 0.004147752188146114
step: 200, loss: 0.024049950763583183
step: 210, loss: 0.0419173426926136
step: 220, loss: 0.02709963545203209
step: 230, loss: 0.03541524335741997
step: 240, loss: 0.0036738927010446787
step: 250, loss: 0.00726058054715395
step: 260, loss: 0.00021591050608549267
step: 270, loss: 0.002362126251682639
step: 280, loss: 0.00043605430983006954
step: 290, loss: 0.0014252223772928119
step: 300, loss: 0.002279561012983322
step: 310, loss: 0.00033067166805267334
step: 320, loss: 0.005799380596727133
step: 330, loss: 0.0016916552558541298
step: 340, loss: 0.019373614341020584
step: 350, loss: 0.00031885987846180797
step: 360, loss: 0.0020292308181524277
step: 370, loss: 0.001596591784618795
step: 380, loss: 0.016753660514950752
step: 390, loss: 0.0001923894160427153
step: 400, loss: 0.0010368662187829614
step: 410, loss: 0.009381980635225773
step: 420, loss: 0.00039900632691569626
step: 430, loss: 0.001508083427324891
step: 440, loss: 0.0015798575477674603
step: 450, loss: 0.11287081986665726
step: 460, loss: 0.006084739230573177
step: 470, loss: 0.003621234092861414
step: 480, loss: 0.06646622717380524


```
step: 490, loss: 0.0007750962395220995
step: 500, loss: 0.0013874927535653114
step: 510, loss: 0.005467647686600685
step: 520, loss: 0.019000083208084106
step: 530, loss: 0.042461808770895004
/usr/local/lib/python3.7/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=0.9225092250922509, f1=0.9195722919572292,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005137793254107237
step: 10, loss: 0.008856149390339851
step: 20, loss: 0.0005093455547466874
step: 30, loss: 0.036518849432468414
step: 40, loss: 0.003328334540128708
step: 50, loss: 0.006448804866522551
step: 60, loss: 0.0011572460643947124
step: 70, loss: 0.0007040445925667882
step: 80, loss: 0.017482995986938477
step: 90, loss: 0.007395931053906679
step: 100, loss: 0.0008542946889065206
step: 110, loss: 0.0011273723794147372
step: 120, loss: 0.03289022296667099
step: 130, loss: 0.0006189353880472481
step: 140, loss: 0.03429494798183441
step: 150, loss: 0.017878137528896332
step: 160, loss: 0.016640251502394676
step: 170, loss: 0.002071958966553211
step: 180, loss: 0.005777048412710428
step: 190, loss: 0.0014480884419754148
step: 200, loss: 0.048318345099687576
step: 210, loss: 0.00029477200587280095
step: 220, loss: 0.0019532025326043367
step: 230, loss: 0.10354685038328171
step: 240, loss: 0.0011498159728944302
step: 250, loss: 0.00013779581058770418
step: 260, loss: 0.00024680010392330587
step: 270, loss: 0.0005440701497718692
step: 280, loss: 0.0064694941975176334
step: 290, loss: 0.001039732713252306
step: 300, loss: 0.2346849888563156
step: 310, loss: 0.004929813090711832
step: 320, loss: 0.036620303988456726
step: 330, loss: 0.001785052358172834
step: 340, loss: 0.003638421418145299
```

```

step: 350, loss: 0.0034470870159566402
step: 360, loss: 0.01074246596544981
step: 370, loss: 0.0002414969785604626
step: 380, loss: 0.0023758255410939455
step: 390, loss: 0.06079281494021416
step: 400, loss: 0.005567567422986031
step: 410, loss: 0.00018764453125186265
step: 420, loss: 0.0005193732213228941
step: 430, loss: 0.02548982761800289
step: 440, loss: 0.0020352185238152742
step: 450, loss: 0.021366208791732788
step: 460, loss: 0.03081520088016987
step: 470, loss: 0.0012490105582401156
step: 480, loss: 0.0008698389865458012
step: 490, loss: 0.004191630054265261
step: 500, loss: 0.00021545946947298944
step: 510, loss: 0.0003143523936159909
step: 520, loss: 0.025214768946170807
step: 530, loss: 0.001965279458090663
/usr/local/lib/python3.7/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=0.9287037037037037, f1=0.92, best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015925336629152298
step: 10, loss: 0.001498356694355607
step: 20, loss: 0.0013642202829942107
step: 30, loss: 0.18214558064937592
step: 40, loss: 0.0006699461955577135
step: 50, loss: 0.003370097605511546
step: 60, loss: 0.00011417576752137393
step: 70, loss: 0.0003650700964499265
step: 80, loss: 0.010658716782927513
step: 90, loss: 0.00014160547289066017
step: 100, loss: 0.00013701750140171498
step: 110, loss: 0.0001887880207505077
step: 120, loss: 0.0167680736631155
step: 130, loss: 0.0015683777164667845
step: 140, loss: 0.03721073269844055
step: 150, loss: 0.0001401908230036497
step: 160, loss: 0.0013843589695170522
step: 170, loss: 0.008580139838159084
step: 180, loss: 0.00017711824330035597
step: 190, loss: 0.0013952481094747782
step: 200, loss: 0.010283044539391994
step: 210, loss: 0.0008991220965981483

```

```

step: 220, loss: 0.2084135115146637
step: 230, loss: 0.0004035864258185029
step: 240, loss: 0.0006633665761910379
step: 250, loss: 0.007751569617539644
step: 260, loss: 0.0008676587603986263
step: 270, loss: 0.0014245775528252125
step: 280, loss: 0.009231457486748695
step: 290, loss: 0.07448497414588928
step: 300, loss: 0.0001320867013419047
step: 310, loss: 0.018353290855884552
step: 320, loss: 0.00020420050714164972
step: 330, loss: 0.0011392900487408042
step: 340, loss: 0.11981268227100372
step: 350, loss: 0.0014758831821382046
step: 360, loss: 0.0003387987962923944
step: 370, loss: 0.00153501913882792
step: 380, loss: 0.00022152371820993721
step: 390, loss: 0.00023914848861750215
step: 400, loss: 0.015989229083061218
step: 410, loss: 0.0005323847872205079
step: 420, loss: 7.537038618465886e-05
step: 430, loss: 0.001007243525236845
step: 440, loss: 0.004895462654531002
step: 450, loss: 0.0028760936111211777
step: 460, loss: 0.002495845314115286
step: 470, loss: 0.00047453283332288265
step: 480, loss: 0.0207693912088871
step: 490, loss: 0.021894704550504684
step: 500, loss: 0.0005167314666323364
step: 510, loss: 0.001988802570849657
step: 520, loss: 0.00016181435785256326
step: 530, loss: 0.000780320493504405
/usr/local/lib/python3.7/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=0.9219330855018587, f1=0.9171374764595104,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023452951572835445
step: 10, loss: 0.00027571863029152155
step: 20, loss: 4.1498093196423724e-05
step: 30, loss: 0.13010597229003906
step: 40, loss: 0.0030397665686905384
step: 50, loss: 5.722797504859045e-05
step: 60, loss: 9.797903476282954e-05
step: 70, loss: 0.016187578439712524

```

```
step: 80, loss: 0.0009814448421820998
step: 90, loss: 0.001039348542690277
step: 100, loss: 0.0004005725495517254
step: 110, loss: 0.00031202961690723896
step: 120, loss: 0.0005782077787443995
step: 130, loss: 0.0002698406751733273
step: 140, loss: 0.0003365534939803183
step: 150, loss: 0.0003713882470037788
step: 160, loss: 0.00044277310371398926
step: 170, loss: 5.429775410448201e-05
step: 180, loss: 0.02156946063041687
step: 190, loss: 0.0027359891682863235
step: 200, loss: 0.0006807338795624673
step: 210, loss: 0.00045684020733460784
step: 220, loss: 0.008114502765238285
step: 230, loss: 0.0010968904243782163
step: 240, loss: 0.0003914592962246388
step: 250, loss: 4.9074973503593355e-05
step: 260, loss: 0.00530039519071579
step: 270, loss: 0.0005754518788307905
step: 280, loss: 9.058792056748644e-05
step: 290, loss: 0.00016192779003176838
step: 300, loss: 9.626334212953225e-05
step: 310, loss: 6.534639396704733e-05
step: 320, loss: 0.01255293469876051
step: 330, loss: 8.838607027428225e-05
step: 340, loss: 0.0007839479949325323
step: 350, loss: 0.0008207810460589826
step: 360, loss: 0.003807318164035678
step: 370, loss: 0.000621844781562686
step: 380, loss: 0.0010481536155566573
step: 390, loss: 0.00047635508235543966
step: 400, loss: 0.0006579991313628852
step: 410, loss: 0.0006791470805183053
step: 420, loss: 0.001956664491444826
step: 430, loss: 0.00020938630041200668
step: 440, loss: 0.0016979749780148268
step: 450, loss: 7.4296876846347e-05
step: 460, loss: 0.003497167956084013
step: 470, loss: 0.0018715716432780027
step: 480, loss: 0.0011237122816964984
step: 490, loss: 0.008027471601963043
step: 500, loss: 0.0025281058624386787
step: 510, loss: 0.003768413793295622
step: 520, loss: 0.00033995506237261
step: 530, loss: 0.0015837561804801226
/usr/local/lib/python3.7/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=0.932093023255814, f1=0.9193473193473194,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009648554027080536
step: 10, loss: 0.0014730826951563358
step: 20, loss: 0.0015645360108464956
step: 30, loss: 0.005137935746461153
step: 40, loss: 0.0007339115836657584
step: 50, loss: 0.016778549179434776
step: 60, loss: 0.020633015781641006
step: 70, loss: 0.0006614789599552751
step: 80, loss: 0.00012834560766350478
step: 90, loss: 7.556869240943342e-05
step: 100, loss: 0.002190470462664962
step: 110, loss: 0.004513822495937347
step: 120, loss: 0.0003352385538164526
step: 130, loss: 0.0007337981951422989
step: 140, loss: 0.03224184736609459
step: 150, loss: 0.00014825454854872078
step: 160, loss: 0.0003263356047682464
step: 170, loss: 0.00015598467143718153
step: 180, loss: 0.0012300768867135048
step: 190, loss: 0.014015035703778267
step: 200, loss: 0.0014311146223917603
step: 210, loss: 0.003826592117547989
step: 220, loss: 5.3250380005920306e-05
step: 230, loss: 0.013083573430776596
step: 240, loss: 0.00019193573098164052
step: 250, loss: 0.0001154731071437709
step: 260, loss: 0.0018619159236550331
step: 270, loss: 0.014798481948673725
step: 280, loss: 0.004033248405903578
step: 290, loss: 0.0005311001441441476
step: 300, loss: 0.0024223732762038708
step: 310, loss: 0.0019212928600609303
step: 320, loss: 2.8103017029934563e-05
step: 330, loss: 0.00024643776123411953
step: 340, loss: 0.0430515855550766
step: 350, loss: 0.020319471135735512
step: 360, loss: 0.0010565229458734393
step: 370, loss: 0.00392299285158515
step: 380, loss: 0.0002386645646765828
step: 390, loss: 0.00021941735758446157
step: 400, loss: 0.00035525060957297683
step: 410, loss: 0.0006681557279080153
```

```

step: 420, loss: 0.0001015421366901137
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 4.491776053328067e-05
step: 440, loss: 0.005648601800203323
step: 450, loss: 0.0007369777886196971
step: 460, loss: 0.09403887391090393
step: 470, loss: 0.0003604585363063961
step: 480, loss: 0.004419274162501097
step: 490, loss: 0.007778267841786146
step: 500, loss: 0.0016860326286405325
step: 510, loss: 0.0005244927015155554
step: 520, loss: 3.531005859258585e-05
step: 530, loss: 3.958831075578928e-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=0.9241507677989763, f1=0.9148737137511693,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.7976744351908565e-05
step: 10, loss: 0.00453378027305007
step: 20, loss: 0.00013647647574543953
step: 30, loss: 0.00025103759253397584
step: 40, loss: 0.0003325236029922962
step: 50, loss: 0.01465064287185669
step: 60, loss: 0.00044825527584180236
step: 70, loss: 0.15982326865196228
step: 80, loss: 0.0005788172711618245
step: 90, loss: 0.00017035521159414202
step: 100, loss: 0.0029780941549688578
step: 110, loss: 0.0010359918233007193
step: 120, loss: 3.872621527989395e-05
step: 130, loss: 0.0006514078704640269
step: 140, loss: 0.00040602320223115385
step: 150, loss: 0.00010274594387738034
step: 160, loss: 0.00011246476060478017
step: 170, loss: 0.0027171999681741
step: 180, loss: 0.0003795503871515393
step: 190, loss: 0.0003799001278821379
step: 200, loss: 0.0015847990289330482
step: 210, loss: 0.00013752849190495908
step: 220, loss: 0.00014290341641753912
step: 230, loss: 4.912991062155925e-05
step: 240, loss: 7.937271584523842e-05
step: 250, loss: 0.00030557456193491817
step: 260, loss: 0.0018788926536217332

```

```

step: 270, loss: 0.0006839706329628825
step: 280, loss: 0.00039404002018272877
step: 290, loss: 0.0030173659324645996
step: 300, loss: 0.001254754839465022
step: 310, loss: 0.013434304855763912
step: 320, loss: 0.0010209701722487807
step: 330, loss: 0.0470103919506073
step: 340, loss: 0.004846748895943165
step: 350, loss: 0.0005735858576372266
step: 360, loss: 0.0016191109316423535
step: 370, loss: 0.0005391429294832051
step: 380, loss: 0.00014057215594220906
step: 390, loss: 8.740303746890277e-05
step: 400, loss: 0.00016609275189694017
step: 410, loss: 0.0013157760258764029
step: 420, loss: 0.030326448380947113
step: 430, loss: 0.0047794533893466
step: 440, loss: 0.0054856459610164165
step: 450, loss: 0.00016748683992773294
step: 460, loss: 0.00027676744502969086
step: 470, loss: 0.0022724654991179705
step: 480, loss: 0.0010316347470507026
step: 490, loss: 0.0003369718906469643
step: 500, loss: 0.00044727270142175257
step: 510, loss: 0.09788010269403458
step: 520, loss: 0.00030363525729626417
step: 530, loss: 6.03373373451177e-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=0.9321016166281756, f1=0.9192200557103064,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003478733997326344
step: 10, loss: 5.62048917345237e-05
step: 20, loss: 0.0002541427093092352
step: 30, loss: 0.0004082628001924604
step: 40, loss: 4.509397695073858e-05
step: 50, loss: 0.0007008818793110549
step: 60, loss: 0.00015468949277419597
step: 70, loss: 6.359146937029436e-05
step: 80, loss: 0.00024346160353161395
step: 90, loss: 0.002202074509114027
step: 100, loss: 0.0010477615287527442
step: 110, loss: 6.761640543118119e-05
step: 120, loss: 0.0005681024631485343

```

```

step: 130, loss: 5.152808080310933e-05
step: 140, loss: 0.0003526423533912748
step: 150, loss: 0.0008371495641767979
step: 160, loss: 1.896144385682419e-05
step: 170, loss: 0.0003968366654589772
step: 180, loss: 3.00009451166261e-05
step: 190, loss: 0.00043309558532200754
step: 200, loss: 4.547261050902307e-05
step: 210, loss: 6.386667519109324e-05
step: 220, loss: 0.0008183420868590474
step: 230, loss: 0.0037525806110352278
step: 240, loss: 0.0004169284366071224
step: 250, loss: 9.021376899909228e-05
step: 260, loss: 0.00015537287981715053
step: 270, loss: 2.890629184548743e-05
step: 280, loss: 6.422486330848187e-05
step: 290, loss: 0.0003069167141802609
step: 300, loss: 0.00010653363278834149
step: 310, loss: 0.00047047759289853275
step: 320, loss: 0.0011517369421198964
step: 330, loss: 0.0007114222389645875
step: 340, loss: 0.0002452573971822858
step: 350, loss: 0.00010331761586712673
step: 360, loss: 0.00012704666005447507
step: 370, loss: 0.00020429643336683512
step: 380, loss: 0.0003626430407166481
step: 390, loss: 0.0001803097693482414
step: 400, loss: 0.00016572322056163102
step: 410, loss: 0.0023712844122201204
step: 420, loss: 0.00024823591229505837
step: 430, loss: 8.453380723949522e-05
step: 440, loss: 0.00010816628491738811
step: 450, loss: 0.00020298543677199632
step: 460, loss: 0.12426058948040009
step: 470, loss: 0.00018528477812651545
step: 480, loss: 0.00019839424930978566
step: 490, loss: 0.00013468347606249154
step: 500, loss: 0.00017055260832421482
step: 510, loss: 0.0006410331116057932
step: 520, loss: 0.0005650052335113287
step: 530, loss: 0.00022023853671271354
/usr/local/lib/python3.7/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=0.9327731092436975, f1=0.9229332087809435,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026153854560106993
step: 10, loss: 4.9310219765175134e-05
step: 20, loss: 0.015701856464147568
step: 30, loss: 0.00010518881754251197
step: 40, loss: 0.000268953968770802
step: 50, loss: 0.0025954940356314182
step: 60, loss: 2.6820462153409608e-05
step: 70, loss: 0.0006740607786923647
step: 80, loss: 1.6223460988840088e-05
step: 90, loss: 0.00015907200577203184
step: 100, loss: 0.000516916043125093
step: 110, loss: 6.808435136917979e-05
step: 120, loss: 0.00040977520984597504
step: 130, loss: 0.007055923342704773
step: 140, loss: 0.006356517784297466
step: 150, loss: 7.079699571477249e-05
step: 160, loss: 0.03247131034731865
step: 170, loss: 0.00014076365914661437
step: 180, loss: 3.760147956199944e-05
step: 190, loss: 0.00044484235695563257
step: 200, loss: 3.7629382859449834e-05
step: 210, loss: 4.9204780225409195e-05
step: 220, loss: 3.5737768484978005e-05
step: 230, loss: 0.0009452075464650989
step: 240, loss: 6.0435457271523774e-05
step: 250, loss: 0.00023151611094363034
step: 260, loss: 0.00014729209942743182
step: 270, loss: 5.098446854390204e-05
step: 280, loss: 5.18830529472325e-05
step: 290, loss: 0.009584562852978706
step: 300, loss: 0.0013774977996945381
step: 310, loss: 0.0034352163784205914
step: 320, loss: 0.00014400806685443968
step: 330, loss: 0.0002819234214257449
step: 340, loss: 6.247832789085805e-05
step: 350, loss: 0.00045235914876684546
step: 360, loss: 0.0274328775703907
step: 370, loss: 7.326358172576874e-05
step: 380, loss: 0.20823337137699127
step: 390, loss: 0.028580276295542717
step: 400, loss: 0.0005316753522492945
step: 410, loss: 9.134408901445568e-05
step: 420, loss: 0.0002018825034610927
step: 430, loss: 0.0077605536207556725
step: 440, loss: 0.0005825881962664425
step: 450, loss: 0.0010047501418739557
step: 460, loss: 0.0006117084994912148
```

```

step: 470, loss: 0.0009035448310896754
step: 480, loss: 1.8436181562719867e-05
step: 490, loss: 6.5370746597182e-05
step: 500, loss: 3.5275912523502484e-05
step: 510, loss: 0.0004940627841278911
step: 520, loss: 8.238494774559513e-05
step: 530, loss: 0.0005817164201289415
/usr/local/lib/python3.7/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=0.9290322580645162, f1=0.9232192414431082,
best_f1=0.9196597353497165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.2226117986720055e-05
step: 10, loss: 4.772039028466679e-05
step: 20, loss: 0.0004965105908922851
step: 30, loss: 5.019307718612254e-05
step: 40, loss: 0.0006072958349250257
step: 50, loss: 0.0002489501202944666
step: 60, loss: 4.254690793459304e-05
step: 70, loss: 0.0003655271721072495
step: 80, loss: 8.726802479941398e-05
step: 90, loss: 3.0013425202923827e-05
step: 100, loss: 0.01048039086163044
step: 110, loss: 0.00013832998229190707
step: 120, loss: 0.00011888363223988563
step: 130, loss: 0.11310571432113647
step: 140, loss: 4.731099761556834e-05
step: 150, loss: 9.170676639769226e-05
step: 160, loss: 4.782206451636739e-05
step: 170, loss: 0.0028793932870030403
step: 180, loss: 0.0009454201790504158
step: 190, loss: 0.00042771216249093413
step: 200, loss: 0.0010804517660290003
step: 210, loss: 0.0009098067530430853
step: 220, loss: 4.089293361175805e-05
step: 230, loss: 0.0003361987473908812
step: 240, loss: 0.0005004301201552153
step: 250, loss: 8.414097828790545e-05
step: 260, loss: 0.0003148524556308985
step: 270, loss: 2.5807876227190718e-05
step: 280, loss: 3.4672753827180713e-05
step: 290, loss: 2.2336233087116852e-05
step: 300, loss: 4.249016274116002e-05
step: 310, loss: 0.000482159317471087
step: 320, loss: 0.0007073202286846936

```

```

step: 330, loss: 3.36774428433273e-05
step: 340, loss: 4.954818359692581e-05
step: 350, loss: 0.00012025028991047293
step: 360, loss: 0.0005635999841615558
step: 370, loss: 0.002148512750864029
step: 380, loss: 7.466197712346911e-05
step: 390, loss: 0.055108413100242615
step: 400, loss: 0.0003677912172861397
step: 410, loss: 0.0008099242695607245
step: 420, loss: 0.0003363520954735577
step: 430, loss: 0.002222346160560846
step: 440, loss: 0.00023725036589894444
step: 450, loss: 2.2686443116981536e-05
step: 460, loss: 0.00019538530614227057
step: 470, loss: 0.031754836440086365
step: 480, loss: 6.020467117195949e-05
step: 490, loss: 2.1624739019898698e-05
step: 500, loss: 0.0007931701838970184
step: 510, loss: 0.00011863540566992015
step: 520, loss: 6.96711431373842e-05
step: 530, loss: 5.3304818720789626e-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=0.9291412482402628, f1=0.9189443920829407,
best_f1=0.9196597353497165
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, 246.14it/s]
load_f1 = 0.9343880874825501
real_f1 = 0.9333958724202628
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 250.42it/s]

```

4.1.4 Amazon-Google - Running the matcher

```

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

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

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/Amazon-Google \
    --input_path input/input_small.jsonl \

```

```

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

```

Downloading: 100% 433/433 [00:00<00:00, 360kB/s]
 Downloading: 100% 232k/232k [00:00<00:00, 256kB/s]
 Downloading: 100% 440M/440M [00:06<00:00, 69.7MB/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: 0.5637221932411194

step: 10, loss: 0.37322118878364563

```

step: 20, loss: 0.4058138430118561
step: 30, loss: 0.3283562660217285
step: 40, loss: 0.18619608879089355
step: 50, loss: 0.4480651021003723
step: 60, loss: 0.3448313772678375
step: 70, loss: 0.22345556318759918
step: 80, loss: 0.1972709596157074
step: 90, loss: 0.44629988074302673
step: 100, loss: 0.5356988310813904
step: 110, loss: 0.2789957821369171
step: 120, loss: 0.2572382092475891
step: 130, loss: 0.2957536280155182
step: 140, loss: 0.2958929240703583
step: 150, loss: 0.26097387075424194
step: 160, loss: 0.4113193452358246
step: 170, loss: 0.321792334318161
step: 180, loss: 0.16032350063323975
step: 190, loss: 0.22467634081840515
step: 200, loss: 0.3120046555995941
step: 210, loss: 0.25046154856681824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5349650349650349, f1=0.508833922261484,
best_f1=0.508833922261484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1527014970779419
step: 10, loss: 0.24980796873569489
step: 20, loss: 0.26450955867767334
step: 30, loss: 0.252880334854126
step: 40, loss: 0.30408361554145813
step: 50, loss: 0.22848939895629883
step: 60, loss: 0.5252320170402527
step: 70, loss: 0.18501169979572296
step: 80, loss: 0.26543229818344116
step: 90, loss: 0.12392950803041458
step: 100, loss: 0.028897304087877274
step: 110, loss: 0.15039411187171936
step: 120, loss: 0.18817487359046936
step: 130, loss: 0.03461533039808273
step: 140, loss: 0.21982714533805847
step: 150, loss: 0.24537746608257294

```

```

step: 160, loss: 0.11724188178777695
step: 170, loss: 0.09750554710626602
step: 180, loss: 0.257241427898407
step: 190, loss: 0.2502919137477875
step: 200, loss: 0.07435597479343414
step: 210, loss: 0.12305959314107895
/usr/local/lib/python3.7/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=0.42760487144790255, f1=0.3053691275167785,
best_f1=0.508833922261484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11933503299951553
step: 10, loss: 0.1824105829000473
step: 20, loss: 0.21026161313056946
step: 30, loss: 0.24825718998908997
step: 40, loss: 0.09000888466835022
step: 50, loss: 0.1551463007926941
step: 60, loss: 0.2101762741804123
step: 70, loss: 0.1291135996580124
step: 80, loss: 0.2196958065032959
step: 90, loss: 0.10433695465326309
step: 100, loss: 0.24151577055454254
step: 110, loss: 0.22242291271686554
step: 120, loss: 0.09595778584480286
step: 130, loss: 0.181850865483284
step: 140, loss: 0.1021953597664833
step: 150, loss: 0.22795280814170837
step: 160, loss: 0.04150907322764397
step: 170, loss: 0.163547083735466
step: 180, loss: 0.09416591376066208
step: 190, loss: 0.2245461344718933
step: 200, loss: 0.10783226788043976
step: 210, loss: 0.07452119141817093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5814432989690721, f1=0.5738396624472575,
best_f1=0.5738396624472575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.16190622746944427
step: 10, loss: 0.20006293058395386
step: 20, loss: 0.04405095800757408
step: 30, loss: 0.26014864444732666
step: 40, loss: 0.07090122997760773
step: 50, loss: 0.24983780086040497
step: 60, loss: 0.1798391044139862
step: 70, loss: 0.3370791971683502
step: 80, loss: 0.2178480625152588
step: 90, loss: 0.04154952988028526
step: 100, loss: 0.24356763064861298
step: 110, loss: 0.14198707044124603
step: 120, loss: 0.10241132229566574
step: 130, loss: 0.10532030463218689
step: 140, loss: 0.3065461814403534
step: 150, loss: 0.12450753152370453
step: 160, loss: 0.054300349205732346
step: 170, loss: 0.09135830402374268
step: 180, loss: 0.4809528887271881
step: 190, loss: 0.05463770404458046
step: 200, loss: 0.08530940860509872
step: 210, loss: 0.23906628787517548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6006944444444444, f1=0.5781818181818182,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16249464452266693
step: 10, loss: 0.04690401628613472
step: 20, loss: 0.1810503602027893
step: 30, loss: 0.06871730089187622
step: 40, loss: 0.08771375566720963
step: 50, loss: 0.08357657492160797
step: 60, loss: 0.03382812812924385
step: 70, loss: 0.1379304975271225
step: 80, loss: 0.07324764877557755
step: 90, loss: 0.170519158244133
step: 100, loss: 0.053833574056625366
step: 110, loss: 0.11426226049661636
step: 120, loss: 0.1485510617494583
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```



```

step: 130, loss: 0.10192819684743881
step: 140, loss: 0.16031907498836517
step: 150, loss: 0.09426049143075943
step: 160, loss: 0.22991780936717987
step: 170, loss: 0.06564833968877792
step: 180, loss: 0.08425232768058777
step: 190, loss: 0.044832102954387665
step: 200, loss: 0.06129204481840134
step: 210, loss: 0.013434928841888905
/usr/local/lib/python3.7/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=0.5451055662188099, f1=0.5749486652977414,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05732279270887375
step: 10, loss: 0.059095196425914764
step: 20, loss: 0.02064846269786358
step: 30, loss: 0.005576029419898987
step: 40, loss: 0.009797834791243076
step: 50, loss: 0.02298280969262123
step: 60, loss: 0.12648709118366241
step: 70, loss: 0.02595781348645687
step: 80, loss: 0.07259595394134521
step: 90, loss: 0.03591737523674965
step: 100, loss: 0.008664164692163467
step: 110, loss: 0.03703608363866806
step: 120, loss: 0.03685236722230911
step: 130, loss: 0.03028428740799427
step: 140, loss: 0.07071328163146973
step: 150, loss: 0.04352534934878349
step: 160, loss: 0.01757989637553692
step: 170, loss: 0.20071540772914886
step: 180, loss: 0.08520490676164627
step: 190, loss: 0.11477433145046234
step: 200, loss: 0.07149648666381836
step: 210, loss: 0.005128505639731884
/usr/local/lib/python3.7/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=0.5493716337522441, f1=0.5549132947976879,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0056623294949531555

```

```

step: 10, loss: 0.009531920775771141
step: 20, loss: 0.1432536542415619
step: 30, loss: 0.027234258130192757
step: 40, loss: 0.10450942814350128
step: 50, loss: 0.13409368693828583
step: 60, loss: 0.03581269830465317
step: 70, loss: 0.025045912712812424
step: 80, loss: 0.11624684929847717
step: 90, loss: 0.03848176449537277
step: 100, loss: 0.015475082211196423
step: 110, loss: 0.01543775200843811
step: 120, loss: 0.2908589541912079
step: 130, loss: 0.050630345940589905
step: 140, loss: 0.013834765180945396
step: 150, loss: 0.012677519582211971
step: 160, loss: 0.06828250735998154
step: 170, loss: 0.04840552806854248
step: 180, loss: 0.037505704909563065
step: 190, loss: 0.26067808270454407
step: 200, loss: 0.038433536887168884
step: 210, loss: 0.08347822725772858
/usr/local/lib/python3.7/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=0.5438898450946643, f1=0.5498154981549815,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02571113407611847
step: 10, loss: 0.14037109911441803
step: 20, loss: 0.01960286870598793
step: 30, loss: 0.011539334431290627
step: 40, loss: 0.033172864466905594
step: 50, loss: 0.011553257703781128
step: 60, loss: 0.17689922451972961
step: 70, loss: 0.21735741198062897
step: 80, loss: 0.10338274389505386
step: 90, loss: 0.015464047901332378
step: 100, loss: 0.1075035110116005
step: 110, loss: 0.03895305469632149
step: 120, loss: 0.015301062725484371
step: 130, loss: 0.1512092500925064
step: 140, loss: 0.044815436005592346
step: 150, loss: 0.013604756444692612
step: 160, loss: 0.06810549646615982
step: 170, loss: 0.12134283035993576
step: 180, loss: 0.08261843025684357

```

```

step: 190, loss: 0.0024103308096528053
step: 200, loss: 0.01068649161607027
step: 210, loss: 0.1642647683620453
/usr/local/lib/python3.7/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=0.5333333333333333, f1=0.5107398568019093,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010578351095318794
step: 10, loss: 0.020468972623348236
step: 20, loss: 0.005798340309411287
step: 30, loss: 0.036626435816287994
step: 40, loss: 0.028985794633626938
step: 50, loss: 0.01222226768732071
step: 60, loss: 0.2687849700450897
step: 70, loss: 0.09309873729944229
step: 80, loss: 0.008612888865172863
step: 90, loss: 0.0070797149091959
step: 100, loss: 0.04418036714196205
step: 110, loss: 0.04257546365261078
step: 120, loss: 0.029207129031419754
step: 130, loss: 0.08452926576137543
step: 140, loss: 0.03493691235780716
step: 150, loss: 0.044510725885629654
step: 160, loss: 0.010357057675719261
step: 170, loss: 0.00213528610765934
step: 180, loss: 0.002220870228484273
step: 190, loss: 0.002139043528586626
step: 200, loss: 0.01836845837533474
step: 210, loss: 0.022920764982700348
/usr/local/lib/python3.7/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=0.5271028037383179, f1=0.544378698224852,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01264736708253622
step: 10, loss: 0.02066384069621563
step: 20, loss: 0.007928894832730293
step: 30, loss: 0.17348124086856842
step: 40, loss: 0.007521606516093016
step: 50, loss: 0.012125026434659958
step: 60, loss: 0.06326941400766373

```

```

step: 70, loss: 0.02951047383248806
step: 80, loss: 0.05359850451350212
step: 90, loss: 0.06243003159761429
step: 100, loss: 0.011701293289661407
step: 110, loss: 0.001201605424284935
step: 120, loss: 0.0024383151903748512
step: 130, loss: 0.03512076288461685
step: 140, loss: 0.029721638187766075
step: 150, loss: 0.010119745507836342
step: 160, loss: 0.01842697523534298
step: 170, loss: 0.027373261749744415
step: 180, loss: 0.1154976561665535
step: 190, loss: 0.15691031515598297
step: 200, loss: 0.015260211192071438
step: 210, loss: 0.14035043120384216
/usr/local/lib/python3.7/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=0.518375241779497, f1=0.5084033613445379,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01956743746995926
step: 10, loss: 0.2667587101459503
step: 20, loss: 0.02683289535343647
step: 30, loss: 0.007700758520513773
step: 40, loss: 0.0778229609131813
step: 50, loss: 0.006963429041206837
step: 60, loss: 0.05221616476774216
step: 70, loss: 0.039674270898103714
step: 80, loss: 0.02629382163286209
step: 90, loss: 0.16162194311618805
step: 100, loss: 0.009773915633559227
step: 110, loss: 0.017953991889953613
step: 120, loss: 0.04514676332473755
step: 130, loss: 0.005479432176798582
step: 140, loss: 0.03344535827636719
step: 150, loss: 0.0005278292228467762
step: 160, loss: 0.015313942916691303
step: 170, loss: 0.0037050337996333838
step: 180, loss: 0.0073637510649859905
step: 190, loss: 0.004490445367991924
step: 200, loss: 0.03265848010778427
step: 210, loss: 0.027508778497576714
/usr/local/lib/python3.7/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=0.5448154657293497, f1=0.5315487571701721,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04898052290081978
step: 10, loss: 0.003996027167886496
step: 20, loss: 0.009557358920574188
step: 30, loss: 0.02078540064394474
step: 40, loss: 0.31218627095222473
step: 50, loss: 0.0006103632622398436
step: 60, loss: 0.003610940184444189
step: 70, loss: 0.016784915700554848
step: 80, loss: 0.048433929681777954
step: 90, loss: 0.14491192996501923
step: 100, loss: 0.00609822291880846
step: 110, loss: 0.007183650974184275
step: 120, loss: 0.0017129338812083006
step: 130, loss: 0.0006662384257651865
step: 140, loss: 0.037150148302316666
step: 150, loss: 0.003379922593012452
step: 160, loss: 0.0031522323843091726
step: 170, loss: 0.01597389206290245
step: 180, loss: 0.07130454480648041
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.1342414915561676
step: 200, loss: 0.0063656894490122795
step: 210, loss: 0.00794886238873005
/usr/local/lib/python3.7/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=0.5211864406779662, f1=0.5136363636363637,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009858491830527782
step: 10, loss: 0.0024870203342288733
step: 20, loss: 0.015070200897753239
step: 30, loss: 0.0332290455698967
step: 40, loss: 0.00294427084736526
step: 50, loss: 0.004404962062835693
step: 60, loss: 0.004718857817351818
step: 70, loss: 0.032279692590236664
step: 80, loss: 0.015477643348276615
step: 90, loss: 0.0014692348195239902
step: 100, loss: 0.004277216270565987
step: 110, loss: 0.002165810205042362

```

```

step: 120, loss: 0.0013343680184334517
step: 130, loss: 0.0086592435836792
step: 140, loss: 0.0040197488851845264
step: 150, loss: 0.0048623401671648026
step: 160, loss: 0.0018030429491773248
step: 170, loss: 0.037113625556230545
step: 180, loss: 0.002650106092914939
step: 190, loss: 0.007818561047315598
step: 200, loss: 0.025233019143342972
step: 210, loss: 0.012563904747366905
/usr/local/lib/python3.7/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=0.5144124168514413, f1=0.48557692307692313,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025614791084080935
step: 10, loss: 0.024341491982340813
step: 20, loss: 0.0026687830686569214
step: 30, loss: 0.03353219851851463
step: 40, loss: 0.0077074444852769375
step: 50, loss: 0.028398148715496063
step: 60, loss: 0.023276833817362785
step: 70, loss: 0.0036435090005397797
step: 80, loss: 0.0033644186332821846
step: 90, loss: 0.006274526473134756
step: 100, loss: 0.0157928504049778
step: 110, loss: 0.009765369817614555
step: 120, loss: 0.029653286561369896
step: 130, loss: 0.18267922103405
step: 140, loss: 0.016931753605604172
step: 150, loss: 0.003915435168892145
step: 160, loss: 0.03396826982498169
step: 170, loss: 0.0009132022969424725
step: 180, loss: 0.11486352235078812
step: 190, loss: 0.007667114958167076
step: 200, loss: 0.05058444291353226
step: 210, loss: 0.020194752141833305
/usr/local/lib/python3.7/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=0.5116279069767442, f1=0.49769585253456217,
best_f1=0.5781818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.002242199145257473
step: 10, loss: 0.0014811642467975616
step: 20, loss: 0.0050668795593082905
step: 30, loss: 0.028698641806840897
step: 40, loss: 0.0038497294299304485
step: 50, loss: 0.002628239570185542
step: 60, loss: 0.005490886978805065
step: 70, loss: 0.0011397798079997301
step: 80, loss: 0.003943422809243202
step: 90, loss: 0.0008395594195462763
step: 100, loss: 0.004124801605939865
step: 110, loss: 0.0010281784925609827
step: 120, loss: 0.003302315017208457
step: 130, loss: 0.14208905398845673
step: 140, loss: 0.0010390948737040162
step: 150, loss: 0.005179142113775015
step: 160, loss: 0.0038387717213481665
step: 170, loss: 0.0014470615424215794
step: 180, loss: 0.00619935430586338
step: 190, loss: 0.13646940886974335
step: 200, loss: 0.004549384117126465
step: 210, loss: 0.004948488436639309
/usr/local/lib/python3.7/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=0.5110132158590308, f1=0.49406175771971494,
best_f1=0.5781818181818182
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 [00:08, 276.42it/s]
load_f1 = 0.578125
real_f1 = 0.5680000000000001
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 189.80it/s]

```

4.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```



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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

```
step: 0, loss: 0.5578705668449402
```

```
step: 10, loss: 0.3626886308193207
```

```
step: 20, loss: 0.2918914258480072
```

```
step: 30, loss: 0.423957496881485
```

```

step: 40, loss: 0.4363842308521271
step: 50, loss: 0.293529212474823
step: 60, loss: 0.2733671963214874
step: 70, loss: 0.2692078948020935
step: 80, loss: 0.2548108994960785
step: 90, loss: 0.28662702441215515
step: 100, loss: 0.292223185300827
step: 110, loss: 0.4509568512439728
step: 120, loss: 0.08724716305732727
step: 130, loss: 0.17084912955760956
step: 140, loss: 0.059136588126420975
step: 150, loss: 0.17632921040058136
step: 160, loss: 0.0806131660938263
step: 170, loss: 0.2581450641155243
step: 180, loss: 0.023794200271368027
step: 190, loss: 0.2480972558259964
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6503667481662592, f1=0.6859903381642513,
best_f1=0.6859903381642513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2989102005958557
step: 10, loss: 0.17310559749603271
step: 20, loss: 0.16326472163200378
step: 30, loss: 0.12110470980405807
step: 40, loss: 0.10634124279022217
step: 50, loss: 0.1762356460094452
step: 60, loss: 0.3164792060852051
step: 70, loss: 0.2983478009700775
step: 80, loss: 0.16236719489097595
step: 90, loss: 0.17102020978927612
step: 100, loss: 0.07458433508872986
step: 110, loss: 0.10392660647630692
step: 120, loss: 0.20600737631320953
step: 130, loss: 0.17433765530586243
step: 140, loss: 0.056982580572366714
step: 150, loss: 0.08697241544723511
step: 160, loss: 0.13516300916671753
step: 170, loss: 0.08540935814380646
step: 180, loss: 0.1088491752743721
step: 190, loss: 0.05565296858549118

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.793388429752066, f1=0.7888888888888889,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052015598863363266
step: 10, loss: 0.11462979018688202
step: 20, loss: 0.13533072173595428
step: 30, loss: 0.08167005330324173
step: 40, loss: 0.12944364547729492
step: 50, loss: 0.29783257842063904
step: 60, loss: 0.038641899824142456
step: 70, loss: 0.06391509622335434
step: 80, loss: 0.12201881408691406
step: 90, loss: 0.06229967623949051
step: 100, loss: 0.08841317892074585
step: 110, loss: 0.03255070373415947
step: 120, loss: 0.08736275136470795
step: 130, loss: 0.01651180163025856
step: 140, loss: 0.102165088057518
step: 150, loss: 0.10196740925312042
step: 160, loss: 0.040555428713560104
step: 170, loss: 0.1330634206533432
step: 180, loss: 0.06385700404644012
step: 190, loss: 0.06047786399722099
/usr/local/lib/python3.7/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=0.7745358090185677, f1=0.8021680216802168,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013468404300510883
step: 10, loss: 0.11972693353891373
step: 20, loss: 0.17753534018993378
step: 30, loss: 0.11367201805114746
step: 40, loss: 0.042877934873104095
step: 50, loss: 0.0480520986020565
step: 60, loss: 0.01635233499109745
step: 70, loss: 0.04827369004487991

```

```

step: 80, loss: 0.06471271067857742
step: 90, loss: 0.017371490597724915
step: 100, loss: 0.09630991518497467
step: 110, loss: 0.009365076199173927
step: 120, loss: 0.20759981870651245
step: 130, loss: 0.29043903946876526
step: 140, loss: 0.023957183584570885
step: 150, loss: 0.03166375681757927
step: 160, loss: 0.02208567038178444
step: 170, loss: 0.12701982259750366
step: 180, loss: 0.031044097617268562
step: 190, loss: 0.31587085127830505
/usr/local/lib/python3.7/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=0.7828418230563003, f1=0.7650273224043715,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07247709482908249
step: 10, loss: 0.01892159692943096
step: 20, loss: 0.02241957187652588
step: 30, loss: 0.057985588908195496
step: 40, loss: 0.019473636522889137
step: 50, loss: 0.2859571874141693
step: 60, loss: 0.07035508006811142
step: 70, loss: 0.026602987200021744
step: 80, loss: 0.023252185434103012
step: 90, loss: 0.004651300143450499
step: 100, loss: 0.004536220338195562
step: 110, loss: 0.0068448088131845
step: 120, loss: 0.008811110630631447
step: 130, loss: 0.21287719905376434
step: 140, loss: 0.03749880939722061
step: 150, loss: 0.1449470967054367
step: 160, loss: 0.041067659854888916
step: 170, loss: 0.00726222712546587
step: 180, loss: 0.09004087746143341
step: 190, loss: 0.04448315501213074
/usr/local/lib/python3.7/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=0.7611111111111111, f1=0.7696629213483147,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.007215600926429033
step: 10, loss: 0.016855573281645775
step: 20, loss: 0.056731294840574265
step: 30, loss: 0.0048003545962274075
step: 40, loss: 0.015663832426071167
step: 50, loss: 0.004230880178511143
step: 60, loss: 0.07877031713724136
step: 70, loss: 0.0029897126369178295
step: 80, loss: 0.012258892878890038
step: 90, loss: 0.08680935204029083
step: 100, loss: 0.0023653353564441204
step: 110, loss: 0.013817904517054558
step: 120, loss: 0.06838928163051605
step: 130, loss: 0.02709972858428955
step: 140, loss: 0.13149961829185486
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.02071191370487213
step: 160, loss: 0.07003176212310791
step: 170, loss: 0.025673460215330124
step: 180, loss: 0.007685062941163778
step: 190, loss: 0.07099471986293793
/usr/local/lib/python3.7/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=0.7650273224043715, f1=0.7700831024930748,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007466086186468601
step: 10, loss: 0.1011352390050888
step: 20, loss: 0.04457789286971092
step: 30, loss: 0.01433361042290926
step: 40, loss: 0.00888463482260704
step: 50, loss: 0.04781978204846382
step: 60, loss: 0.006655238568782806
step: 70, loss: 0.014070795848965645
step: 80, loss: 0.01923707127571106
step: 90, loss: 0.051567740738391876
step: 100, loss: 0.0014540436677634716
step: 110, loss: 0.011773074977099895
step: 120, loss: 0.009915122762322426
step: 130, loss: 0.007212111726403236
step: 140, loss: 0.028954461216926575
step: 150, loss: 0.02360769733786583
step: 160, loss: 0.04507992044091225
step: 170, loss: 0.04336080700159073
step: 180, loss: 0.0016150305746123195

```

```

step: 190, loss: 0.0026168611366301775
/usr/local/lib/python3.7/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=0.7775061124694377, f1=0.7570332480818414,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00128694879822433
step: 10, loss: 0.01776500977575779
step: 20, loss: 0.0030068468768149614
step: 30, loss: 0.1280817985534668
step: 40, loss: 0.0679362565279007
step: 50, loss: 0.03677138686180115
step: 60, loss: 0.0014580771094188094
step: 70, loss: 0.007611452601850033
step: 80, loss: 0.0011104652658104897
step: 90, loss: 0.01028315257281065
step: 100, loss: 0.06309832632541656
step: 110, loss: 0.003884280100464821
step: 120, loss: 0.002481113886460662
step: 130, loss: 0.000879135390277952
step: 140, loss: 0.002536626998335123
step: 150, loss: 0.008253700099885464
step: 160, loss: 0.0033919336274266243
step: 170, loss: 0.002738177077844739
step: 180, loss: 0.02906901016831398
step: 190, loss: 0.13073629140853882
/usr/local/lib/python3.7/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=0.7867036011080332, f1=0.7727272727272727,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001411536242812872
step: 10, loss: 0.009188489057123661
step: 20, loss: 0.02070286124944687
step: 30, loss: 0.0013727534096688032
step: 40, loss: 0.005530874244868755
step: 50, loss: 0.06956753134727478
step: 60, loss: 0.0004995325580239296
step: 70, loss: 0.0022837116848677397
step: 80, loss: 0.0018710348522290587
step: 90, loss: 0.0010210657492280006
step: 100, loss: 0.014754966832697392

```

```

step: 110, loss: 0.005504991393536329
step: 120, loss: 0.004573010839521885
step: 130, loss: 0.005151443649083376
step: 140, loss: 0.01077155489474535
step: 150, loss: 0.0017572811339050531
step: 160, loss: 0.052048537880182266
step: 170, loss: 0.0004990879679098725
step: 180, loss: 0.13726851344108582
step: 190, loss: 0.038217395544052124
/usr/local/lib/python3.7/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=0.768421052631579, f1=0.7667560321715818,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00446931179612875
step: 10, loss: 0.024405675008893013
step: 20, loss: 0.016372837126255035
step: 30, loss: 0.001969308126717806
step: 40, loss: 0.01556107122451067
step: 50, loss: 0.027681104838848114
step: 60, loss: 0.0010508125415071845
step: 70, loss: 0.0023343744687736034
step: 80, loss: 0.04352743551135063
step: 90, loss: 0.0011101510608568788
step: 100, loss: 0.005948272533714771
step: 110, loss: 0.014447400346398354
step: 120, loss: 0.0659952163696289
step: 130, loss: 0.03055892512202263
step: 140, loss: 0.0005329878185875714
step: 150, loss: 0.00743176368996501
step: 160, loss: 0.00946818944066763
step: 170, loss: 0.0003318750823382288
step: 180, loss: 0.00025518686743453145
step: 190, loss: 0.00033794049522839487
/usr/local/lib/python3.7/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=0.7872340425531915, f1=0.7563025210084034,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001191307557746768
step: 10, loss: 0.0007061935029923916
step: 20, loss: 0.00831205677241087

```

```

step: 30, loss: 0.03707578778266907
step: 40, loss: 0.00021118215227033943
step: 50, loss: 0.002727067330852151
step: 60, loss: 0.0017391765723004937
step: 70, loss: 0.0004465749370865524
step: 80, loss: 0.0004933099262416363
step: 90, loss: 0.00027276555192656815
step: 100, loss: 0.0002987127227243036
step: 110, loss: 0.00034222027170471847
step: 120, loss: 0.1693909615278244
step: 130, loss: 0.0006502164760604501
step: 140, loss: 0.002745754551142454
step: 150, loss: 0.0007176757790148258
step: 160, loss: 0.0011622633319348097
step: 170, loss: 0.000957400188781321
step: 180, loss: 0.008139640092849731
step: 190, loss: 0.002378226490691304
/usr/local/lib/python3.7/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=0.7925531914893618, f1=0.7808988764044944,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007759183063171804
step: 10, loss: 0.0032890397123992443
step: 20, loss: 0.0003453049575909972
step: 30, loss: 0.0036124633625149727
step: 40, loss: 0.003469867166131735
step: 50, loss: 0.0009484647307544947
step: 60, loss: 0.00036373749026097357
step: 70, loss: 0.0006049740477465093
step: 80, loss: 0.0005453290650621057
step: 90, loss: 0.04651809111237526
step: 100, loss: 0.000690993620082736
step: 110, loss: 0.004927595611661673
step: 120, loss: 0.0006806678720749915
step: 130, loss: 0.01910972036421299
step: 140, loss: 0.27138200402259827
step: 150, loss: 0.0016590299783274531
step: 160, loss: 0.0018085670890286565
step: 170, loss: 0.0017795827006921172
step: 180, loss: 0.001452240045182407
step: 190, loss: 0.0012252042070031166
/usr/local/lib/python3.7/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=0.7901907356948229, f1=0.7774647887323942,
best_f1=0.7888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029225563630461693
step: 10, loss: 0.0013465732336044312
step: 20, loss: 0.001101428410038352
step: 30, loss: 0.004964722786098719
step: 40, loss: 0.00423506461083889
step: 50, loss: 0.002282925182953477
step: 60, loss: 0.0006680930964648724
step: 70, loss: 0.09840308874845505
step: 80, loss: 0.0015955906128510833
step: 90, loss: 0.011765284463763237
step: 100, loss: 0.05903543904423714
step: 110, loss: 0.0012283759424462914
step: 120, loss: 0.011822676286101341
step: 130, loss: 0.0007271370268426836
step: 140, loss: 0.0006099938764236867
step: 150, loss: 0.0006504437769763172
step: 160, loss: 0.0006339249666780233
step: 170, loss: 0.00045021346886642277
step: 180, loss: 0.000511994818225503
step: 190, loss: 0.0032776223961263895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7969151670951158, f1=0.7598944591029023,
best_f1=0.7598944591029023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015428734477609396
step: 10, loss: 0.016126738861203194
step: 20, loss: 0.000630234251730144
step: 30, loss: 0.0004625757865142077
step: 40, loss: 0.004007022827863693
step: 50, loss: 0.0007730010547675192
step: 60, loss: 0.03050548955798149
step: 70, loss: 0.0008378959610126913
step: 80, loss: 0.0006760723190382123
step: 90, loss: 0.01925579085946083
step: 100, loss: 0.005820562597364187

```

```

step: 110, loss: 0.0008623694884590805
step: 120, loss: 0.005550340749323368
step: 130, loss: 0.0015486555639654398
step: 140, loss: 0.002378700766712427
step: 150, loss: 0.0007478328188881278
step: 160, loss: 0.0009093750850297511
step: 170, loss: 0.00821592565625906
step: 180, loss: 0.0039994604885578156
step: 190, loss: 0.0004914868623018265
/usr/local/lib/python3.7/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=0.7926509186351706, f1=0.7704918032786885,
best_f1=0.7598944591029023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026396254543215036
step: 10, loss: 0.0030608605593442917
step: 20, loss: 0.010167629458010197
step: 30, loss: 0.0015291128074750304
step: 40, loss: 0.0003924295597244054
step: 50, loss: 0.0036680635530501604
step: 60, loss: 0.0006890893564559519
step: 70, loss: 0.0035725124180316925
step: 80, loss: 0.042386744171381
step: 90, loss: 0.0005874826456420124
step: 100, loss: 0.0007048676488921046
step: 110, loss: 0.0005659778835251927
step: 120, loss: 0.0006032017408870161
step: 130, loss: 0.007743090856820345
step: 140, loss: 0.0030939772259444
step: 150, loss: 0.0008156877593137324
step: 160, loss: 0.0004936092300340533
step: 170, loss: 0.0006796094821766019
step: 180, loss: 0.0007117660716176033
step: 190, loss: 0.0003870003274641931
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7937336814621411, f1=0.7696476964769648,
best_f1=0.7598944591029023
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled          : True

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                  : True
opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 165.70it/s]
load_f1 = 0.7272727272727273
real_f1 = 0.7039106145251397
0it [00:00, ?it/s]/usr/local/lib/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, 186.81it/s]

```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

```

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

```

```

--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.625952422618866
step: 10, loss: 0.3525077998638153
step: 20, loss: 0.3081722557544708
step: 30, loss: 0.40851083397865295
step: 40, loss: 0.26993030309677124
step: 50, loss: 0.255832701921463
step: 60, loss: 0.2669130563735962
step: 70, loss: 0.3694569766521454
step: 80, loss: 0.3816753029823303
step: 90, loss: 0.25418758392333984
step: 100, loss: 0.16813404858112335
step: 110, loss: 0.2823096513748169
step: 120, loss: 0.23089131712913513
step: 130, loss: 0.06289006769657135
step: 140, loss: 0.2285923808813095
step: 150, loss: 0.29912811517715454
step: 160, loss: 0.12859368324279785
step: 170, loss: 0.21285668015480042
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6363636363636362, f1=0.5841584158415842,
best_f1=0.5841584158415842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20652766525745392
step: 10, loss: 0.24510936439037323
step: 20, loss: 0.06209194287657738
step: 30, loss: 0.22885601222515106
step: 40, loss: 0.0840405523777008

```

```

step: 50, loss: 0.17719967663288116
step: 60, loss: 0.16996124386787415
step: 70, loss: 0.09808365255594254
step: 80, loss: 0.10822555422782898
step: 90, loss: 0.09577227383852005
step: 100, loss: 0.1415538191795349
step: 110, loss: 0.09757373481988907
step: 120, loss: 0.11042658984661102
step: 130, loss: 0.09381739050149918
step: 140, loss: 0.3375532627105713
step: 150, loss: 0.34607258439064026
step: 160, loss: 0.20250897109508514
step: 170, loss: 0.07640833407640457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7549668874172185, f1=0.7352297592997812,
best_f1=0.7352297592997812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11632613092660904
step: 10, loss: 0.18427543342113495
step: 20, loss: 0.061824627220630646
step: 30, loss: 0.13709597289562225
step: 40, loss: 0.0848117247223854
step: 50, loss: 0.12597812712192535
step: 60, loss: 0.09690459072589874
step: 70, loss: 0.08103996515274048
step: 80, loss: 0.07157455384731293
step: 90, loss: 0.058988600969314575
step: 100, loss: 0.01646432653069496
step: 110, loss: 0.05436885729432106
step: 120, loss: 0.050331201404333115
step: 130, loss: 0.14294113218784332
step: 140, loss: 0.08919933438301086
step: 150, loss: 0.03297492489218712
step: 160, loss: 0.039994094520807266
step: 170, loss: 0.04046163335442543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7560321715817695, f1=0.7667560321715817,
best_f1=0.7667560321715817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01522791851311922
step: 10, loss: 0.140489861369133
step: 20, loss: 0.032628145068883896
step: 30, loss: 0.1583559662103653
step: 40, loss: 0.006533645559102297
step: 50, loss: 0.07561572641134262
step: 60, loss: 0.061024922877550125
step: 70, loss: 0.007068977225571871
step: 80, loss: 0.07102084159851074
step: 90, loss: 0.11237280070781708
step: 100, loss: 0.15957073867321014
step: 110, loss: 0.03350912034511566
step: 120, loss: 0.005962822586297989
step: 130, loss: 0.1373337060213089
step: 140, loss: 0.04650013521313667
step: 150, loss: 0.13332656025886536
step: 160, loss: 0.10330317914485931
step: 170, loss: 0.014980481937527657
/usr/local/lib/python3.7/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=0.7506053268765133, f1=0.7688564476885645,
best_f1=0.7667560321715817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033184297382831573
step: 10, loss: 0.026305703446269035
step: 20, loss: 0.020508352667093277
step: 30, loss: 0.060272205621004105
step: 40, loss: 0.026357006281614304
step: 50, loss: 0.15411075949668884
step: 60, loss: 0.03692694753408432
step: 70, loss: 0.23285754024982452
step: 80, loss: 0.05302987992763519
step: 90, loss: 0.15464045107364655
step: 100, loss: 0.04379749670624733
step: 110, loss: 0.11298771947622299
step: 120, loss: 0.04739055782556534
step: 130, loss: 0.10006106644868851
step: 140, loss: 0.009466633200645447

```

```

step: 150, loss: 0.08869410306215286
step: 160, loss: 0.06424447149038315
step: 170, loss: 0.06126176193356514
/usr/local/lib/python3.7/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=0.751131221719457, f1=0.7462039045553146,
best_f1=0.7667560321715817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01747114770114422
step: 10, loss: 0.11479255557060242
step: 20, loss: 0.013050973415374756
step: 30, loss: 0.04280718415975571
step: 40, loss: 0.028747539967298508
step: 50, loss: 0.08304421603679657
step: 60, loss: 0.05958903953433037
step: 70, loss: 0.057346176356077194
step: 80, loss: 0.0564996600151062
step: 90, loss: 0.038083404302597046
step: 100, loss: 0.01103364396840334
step: 110, loss: 0.0012466153129935265
step: 120, loss: 0.06820517033338547
step: 130, loss: 0.04313528910279274
step: 140, loss: 0.019594401121139526
step: 150, loss: 0.0464937761425972
step: 160, loss: 0.07833119481801987
step: 170, loss: 0.01642022654414177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7611940298507462, f1=0.7446300715990454,
best_f1=0.7446300715990454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010403214022517204
step: 10, loss: 0.0010692339856177568
step: 20, loss: 0.001176071702502668
step: 30, loss: 0.006431303918361664
step: 40, loss: 0.007647461257874966
step: 50, loss: 0.012561806477606297
step: 60, loss: 0.003703808644786477

```



```

step: 70, loss: 0.003088919445872307
step: 80, loss: 0.062482740730047226
step: 90, loss: 0.0010422248160466552
step: 100, loss: 0.030415233224630356
step: 110, loss: 0.09921933710575104
step: 120, loss: 0.011562615633010864
step: 130, loss: 0.14225582778453827
step: 140, loss: 0.0047186738811433315
step: 150, loss: 0.0038567238952964544
step: 160, loss: 0.0029173309449106455
step: 170, loss: 0.1360301375389099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.761904761904762, f1=0.756476683937824,
best_f1=0.756476683937824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005519278813153505
step: 10, loss: 0.012836702167987823
step: 20, loss: 0.0011446200078353286
step: 30, loss: 0.00851823017001152
step: 40, loss: 0.0007464701193384826
step: 50, loss: 0.0025207228027284145
step: 60, loss: 0.002674876246601343
step: 70, loss: 0.00487583177164197
step: 80, loss: 0.009524631313979626
step: 90, loss: 0.012718945741653442
step: 100, loss: 0.03211895748972893
step: 110, loss: 0.06473487615585327
step: 120, loss: 0.00677892193198204
step: 130, loss: 0.006041001062840223
step: 140, loss: 0.01210788730531931
step: 150, loss: 0.05239735171198845
step: 160, loss: 0.010600057430565357
step: 170, loss: 0.002843456109985709
/usr/local/lib/python3.7/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=0.752577319587629, f1=0.7565011820330969,
best_f1=0.756476683937824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0044074589386582375
step: 10, loss: 0.022622134536504745
step: 20, loss: 0.0007234467775560915
step: 30, loss: 0.003379569621756673
step: 40, loss: 0.0077577996999025345
step: 50, loss: 0.0010105500696226954
step: 60, loss: 0.11459530889987946
step: 70, loss: 0.024590160697698593
step: 80, loss: 0.018678514286875725
step: 90, loss: 0.13583897054195404
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.001042234362103045
step: 110, loss: 0.009012442082166672
step: 120, loss: 0.015268313698470592
step: 130, loss: 0.06395083665847778
step: 140, loss: 0.0023879376240074635
step: 150, loss: 0.010464640334248543
step: 160, loss: 0.07064429670572281
step: 170, loss: 0.014933018945157528
/usr/local/lib/python3.7/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=0.7540983606557378, f1=0.7873303167420815,
best_f1=0.756476683937824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027627434581518173
step: 10, loss: 0.0005469998577609658
step: 20, loss: 0.023025013506412506
step: 30, loss: 0.005220259539783001
step: 40, loss: 0.12451934069395065
step: 50, loss: 0.054982636123895645
step: 60, loss: 0.00036795149208046496
step: 70, loss: 0.009578879922628403
step: 80, loss: 0.0019351819064468145
step: 90, loss: 0.007961640134453773
step: 100, loss: 0.00044725043699145317
step: 110, loss: 0.024178320541977882
step: 120, loss: 0.019106272608041763
step: 130, loss: 0.006593373604118824
step: 140, loss: 0.17833185195922852
step: 150, loss: 0.009862015023827553
step: 160, loss: 0.0023150271736085415
step: 170, loss: 0.0008511741762049496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7675675675675676, f1=0.7621483375959078,
best_f1=0.7621483375959078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043721262365579605
step: 10, loss: 0.010377267375588417
step: 20, loss: 0.0018259717617183924
step: 30, loss: 0.00040087581146508455
step: 40, loss: 0.019399970769882202
step: 50, loss: 0.03470487892627716
step: 60, loss: 0.06812357157468796
step: 70, loss: 0.001973355421796441
step: 80, loss: 0.0006418147240765393
step: 90, loss: 0.0013603006955236197
step: 100, loss: 0.0023134597577154636
step: 110, loss: 0.01676071807742119
step: 120, loss: 0.0007223619031719863
step: 130, loss: 0.0008522393181920052
step: 140, loss: 0.0010095929028466344
step: 150, loss: 0.004855559207499027
step: 160, loss: 0.01072019524872303
step: 170, loss: 0.012756302021443844
/usr/local/lib/python3.7/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=0.7625899280575539, f1=0.7780429594272077,
best_f1=0.7621483375959078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017546273302286863
step: 10, loss: 0.0074477046728134155
step: 20, loss: 0.003684617578983307
step: 30, loss: 0.029736805707216263
step: 40, loss: 0.0009806286543607712
step: 50, loss: 0.0025984495878219604
step: 60, loss: 0.002418078249320388
step: 70, loss: 0.06145671382546425
step: 80, loss: 0.00020708861120510846
step: 90, loss: 0.0010659429244697094
step: 100, loss: 0.003189616836607456
step: 110, loss: 0.023010417819023132

```

```

step: 120, loss: 0.012091847136616707
step: 130, loss: 0.023646818473935127
step: 140, loss: 0.006729762069880962
step: 150, loss: 0.0015474206302314997
step: 160, loss: 0.0002488166792318225
step: 170, loss: 0.00018666185496840626
/usr/local/lib/python3.7/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=0.754257907542579, f1=0.7846889952153109,
best_f1=0.7621483375959078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031186724081635475
step: 10, loss: 0.004311359953135252
step: 20, loss: 0.0021612781565636396
step: 30, loss: 0.0005982537986710668
step: 40, loss: 0.017396438866853714
step: 50, loss: 0.00042631191899999976
step: 60, loss: 0.013084271922707558
step: 70, loss: 0.0006994488649070263
step: 80, loss: 0.00813267007470131
step: 90, loss: 0.00036242674104869366
step: 100, loss: 0.12463638931512833
step: 110, loss: 0.0002800559741444886
step: 120, loss: 0.02598942443728447
step: 130, loss: 0.0010224670404568315
step: 140, loss: 0.0006079206359572709
step: 150, loss: 0.03166602551937103
step: 160, loss: 0.003553350456058979
step: 170, loss: 0.00530720129609108
/usr/local/lib/python3.7/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=0.7595628415300546, f1=0.7708333333333335,
best_f1=0.7621483375959078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018751053139567375
step: 10, loss: 0.00019886715745087713
step: 20, loss: 0.0006040032021701336
step: 30, loss: 0.0006036520935595036
step: 40, loss: 0.0002691537665668875
step: 50, loss: 0.0013680941192433238
step: 60, loss: 0.006754446309059858
step: 70, loss: 0.011237474158406258

```

```

step: 80, loss: 0.000982264638878405
step: 90, loss: 0.0005778914783149958
step: 100, loss: 0.003275736700743437
step: 110, loss: 0.0004260878195054829
step: 120, loss: 0.02459586411714554
step: 130, loss: 0.0028273339848965406
step: 140, loss: 0.00043977733002975583
step: 150, loss: 0.00536982249468565
step: 160, loss: 0.0009550228132866323
step: 170, loss: 0.005422309972345829
/usr/local/lib/python3.7/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=0.7520435967302452, f1=0.7589743589743588,
best_f1=0.7621483375959078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031451418180949986
step: 10, loss: 0.0005856521311216056
step: 20, loss: 0.29986706376075745
step: 30, loss: 0.0005713905557058752
step: 40, loss: 0.00279808952473104
step: 50, loss: 0.0005444561247713864
step: 60, loss: 0.018659032881259918
step: 70, loss: 0.0007899126503616571
step: 80, loss: 0.013428151607513428
step: 90, loss: 0.0006504246848635375
step: 100, loss: 0.0022969357669353485
step: 110, loss: 0.0003343209682498127
step: 120, loss: 0.001497334917075932
step: 130, loss: 0.0012503905454650521
step: 140, loss: 0.0056090583093464375
step: 150, loss: 0.0003889946674462408
step: 160, loss: 0.0034533971920609474
step: 170, loss: 0.004551015794277191
/usr/local/lib/python3.7/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=0.7613941018766757, f1=0.7669172932330827,
best_f1=0.7621483375959078
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, 235.49it/s]
load_f1 = 0.5991379310344827
real_f1 = 0.5810526315789474
0it [00:00, ?it/s]/usr/local/lib/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, 188.07it/s]

```

4.3 DITTO DIRTY

4.3.1 DBLP-ACM - Running the matcher

```

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

```

```

--lm bert-base-uncased \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6057124733924866
step: 10, loss: 0.6213995218276978
step: 20, loss: 0.47980600595474243
step: 30, loss: 0.3148536682128906
step: 40, loss: 0.12738239765167236
step: 50, loss: 0.17168226838111877
step: 60, loss: 0.1583525836467743
step: 70, loss: 0.2725420892238617
step: 80, loss: 0.11313168704509735
step: 90, loss: 0.08261208236217499
step: 100, loss: 0.009592236950993538
step: 110, loss: 0.14167900383472443
step: 120, loss: 0.01632988080382347
step: 130, loss: 0.04305301234126091
step: 140, loss: 0.0028965268284082413
step: 150, loss: 0.025975756347179413
step: 160, loss: 0.01984996162354946
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.12209507822990417
step: 180, loss: 0.11044856160879135
step: 190, loss: 0.07566069811582565
step: 200, loss: 0.1855713576078415
step: 210, loss: 0.011035350151360035
step: 220, loss: 0.00678821699693799
step: 230, loss: 0.055441513657569885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9679558011049725, f1=0.9623059866962307,
best_f1=0.9623059866962307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02110324427485466
step: 10, loss: 0.010361159220337868
step: 20, loss: 0.14997464418411255
step: 30, loss: 0.3194913864135742
step: 40, loss: 0.12295226007699966
step: 50, loss: 0.005621647462248802
step: 60, loss: 0.013333559967577457
step: 70, loss: 0.023809928447008133
step: 80, loss: 0.08957292139530182
step: 90, loss: 0.025708520784974098
step: 100, loss: 0.03779658302664757
step: 110, loss: 0.10496018826961517
step: 120, loss: 0.1450348049402237
step: 130, loss: 0.060332927852869034
step: 140, loss: 0.010311614722013474
step: 150, loss: 0.05974225327372551
step: 160, loss: 0.05121485888957977
step: 170, loss: 0.01228796411305666
step: 180, loss: 0.01603192463517189
step: 190, loss: 0.009538988582789898
step: 200, loss: 0.004213134758174419
step: 210, loss: 0.0034901665057986975
step: 220, loss: 0.11151998490095139
step: 230, loss: 0.07870793342590332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9720044792833147, f1=0.9671574178935448,
best_f1=0.9671574178935448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028665756806731224
step: 10, loss: 0.08778548985719681
step: 20, loss: 0.003302744822576642
step: 30, loss: 0.0975019782781601
step: 40, loss: 0.018366297706961632
step: 50, loss: 0.054736632853746414
step: 60, loss: 0.018170690163969994
step: 70, loss: 0.0033986943308264017
step: 80, loss: 0.010327725671231747
step: 90, loss: 0.09356635808944702
step: 100, loss: 0.003550472669303417

```

```

step: 110, loss: 0.0028747334145009518
step: 120, loss: 0.005735862068831921
step: 130, loss: 0.004853105638176203
step: 140, loss: 0.07607465237379074
step: 150, loss: 0.024394793435931206
step: 160, loss: 0.11229600012302399
step: 170, loss: 0.006763726472854614
step: 180, loss: 0.005883134435862303
step: 190, loss: 0.007203506771475077
step: 200, loss: 0.035413265228271484
step: 210, loss: 0.009617460891604424
step: 220, loss: 0.0007627841550856829
step: 230, loss: 0.0030910803470760584
/usr/local/lib/python3.7/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=0.9642058165548099, f1=0.9675977653631285,
best_f1=0.9671574178935448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014101608656346798
step: 10, loss: 0.0012695161858573556
step: 20, loss: 0.004447423852980137
step: 30, loss: 0.0006431547226384282
step: 40, loss: 0.011638863012194633
step: 50, loss: 0.0008547783363610506
step: 60, loss: 0.06179437041282654
step: 70, loss: 0.0022250525653362274
step: 80, loss: 0.010345589369535446
step: 90, loss: 0.04877012223005295
step: 100, loss: 0.007492808159440756
step: 110, loss: 0.0008158021373674273
step: 120, loss: 0.007405778858810663
step: 130, loss: 0.01434978935867548
step: 140, loss: 0.0007227157475426793
step: 150, loss: 0.09533227980136871
step: 160, loss: 0.026413744315505028
step: 170, loss: 0.008557508699595928
step: 180, loss: 0.0012398118851706386
step: 190, loss: 0.0013607327127829194
step: 200, loss: 0.0018006351310759783
step: 210, loss: 0.007544510532170534
step: 220, loss: 0.002948821522295475
step: 230, loss: 0.011524368077516556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9721913236929923, f1=0.9721913236929923,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003926222852896899
step: 10, loss: 0.0010724078165367246
step: 20, loss: 0.0041624391451478004
step: 30, loss: 0.0011734392028301954
step: 40, loss: 0.0004465973179321736
step: 50, loss: 0.005712466314435005
step: 60, loss: 0.22461213171482086
step: 70, loss: 0.0076493071392178535
step: 80, loss: 0.012803226709365845
step: 90, loss: 0.002559965243563056
step: 100, loss: 0.00173371157143265
step: 110, loss: 0.002472029533237219
step: 120, loss: 0.001225626328960061
step: 130, loss: 0.01187058538198471
step: 140, loss: 0.01409648172557354
step: 150, loss: 0.0005430909222923219
step: 160, loss: 0.01724657602608204
step: 170, loss: 0.002295693149790168
step: 180, loss: 0.003567872801795602
step: 190, loss: 0.14173558354377747
step: 200, loss: 0.05804702267050743
step: 210, loss: 0.013083502650260925
step: 220, loss: 0.0008180678705684841
step: 230, loss: 0.0009880948346108198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9732739420935412, f1=0.9698324022346367,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07404782623052597
step: 10, loss: 0.0015756837092339993
step: 20, loss: 0.0068436008878052235

```

```

step: 30, loss: 0.0007310162181966007
step: 40, loss: 0.00523024657741189
step: 50, loss: 0.011641358025372028
step: 60, loss: 0.000872800825163722
step: 70, loss: 0.004547488410025835
step: 80, loss: 0.001144124660640955
step: 90, loss: 0.006699332036077976
step: 100, loss: 0.06360995024442673
step: 110, loss: 0.011530623771250248
step: 120, loss: 0.0002696417213883251
step: 130, loss: 0.0004971081507392228
step: 140, loss: 0.003967678640037775
step: 150, loss: 0.11112773418426514
step: 160, loss: 0.002262207679450512
step: 170, loss: 0.0013436308363452554
step: 180, loss: 0.006791062653064728
step: 190, loss: 0.04267212375998497
step: 200, loss: 0.059682589024305344
step: 210, loss: 0.001760099083185196
step: 220, loss: 0.000282669992884621
step: 230, loss: 0.03688819706439972
/usr/local/lib/python3.7/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=0.9720044792833147, f1=0.971815107102593,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0040527042001485825
step: 10, loss: 0.00019307414186187088
step: 20, loss: 0.0006653942400589585
step: 30, loss: 0.0005433790502138436
step: 40, loss: 0.0007574761984869838
step: 50, loss: 0.0026349290274083614
step: 60, loss: 0.0006490307860076427
step: 70, loss: 0.0020961256232112646
step: 80, loss: 0.00029193374211899936
step: 90, loss: 0.00010456674499437213
step: 100, loss: 0.00013659520482178777
step: 110, loss: 0.00017782168288249522
step: 120, loss: 0.00014115656085778028
step: 130, loss: 0.0009929328225553036
step: 140, loss: 0.0065470729023218155
step: 150, loss: 0.002216788474470377
step: 160, loss: 0.02143220603466034
step: 170, loss: 0.06436430662870407
step: 180, loss: 0.0013483710354194045

```

```

step: 190, loss: 0.0001906918187160045
step: 200, loss: 0.06852518022060394
step: 210, loss: 0.00015012362564448267
step: 220, loss: 0.0009835275122895837
step: 230, loss: 0.021349117159843445
/usr/local/lib/python3.7/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=0.9698324022346367, f1=0.9698324022346367,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023513850464951247
step: 10, loss: 0.0006722036632709205
step: 20, loss: 0.0007067348924465477
step: 30, loss: 0.00010731792281148955
step: 40, loss: 0.00419318163767457
step: 50, loss: 0.007748474832624197
step: 60, loss: 0.0002706072118598968
step: 70, loss: 0.022710097953677177
step: 80, loss: 0.0011949960608035326
step: 90, loss: 0.00011926361185032874
step: 100, loss: 0.0005797228659503162
step: 110, loss: 0.09175854921340942
step: 120, loss: 0.010422980412840843
step: 130, loss: 0.00010873840074054897
step: 140, loss: 0.00011817735503427684
step: 150, loss: 0.00028246542206034064
step: 160, loss: 0.0005843051476404071
step: 170, loss: 0.0012152550043538213
step: 180, loss: 0.005317084025591612
step: 190, loss: 0.00022385219926945865
step: 200, loss: 0.0011563082225620747
step: 210, loss: 0.0004991605528630316
step: 220, loss: 0.0003030687221325934
step: 230, loss: 0.0007788733346387744
/usr/local/lib/python3.7/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=0.9710467706013363, f1=0.9719416386083053,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041695317486301064
step: 10, loss: 0.10379388183355331
step: 20, loss: 0.022683624178171158

```

```

step: 30, loss: 0.0005402321694418788
step: 40, loss: 0.15195195376873016
step: 50, loss: 0.00020204995234962553
step: 60, loss: 0.000749266124330461
step: 70, loss: 0.013956728391349316
step: 80, loss: 0.0005466856528073549
step: 90, loss: 0.005661474075168371
step: 100, loss: 0.00017432647291570902
step: 110, loss: 7.930080028017983e-05
step: 120, loss: 6.414103700080886e-05
step: 130, loss: 3.98217816837132e-05
step: 140, loss: 5.426473217085004e-05
step: 150, loss: 0.00564982183277607
step: 160, loss: 8.602948219049722e-05
step: 170, loss: 0.00012657819024752825
step: 180, loss: 0.0008134147501550615
step: 190, loss: 0.002412275178357959
step: 200, loss: 0.03639426827430725
step: 210, loss: 0.0003476931597106159
step: 220, loss: 0.0016725059831514955
step: 230, loss: 0.0001010530540952459
/usr/local/lib/python3.7/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=0.9648127128263336, f1=0.9683257918552037,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013529477291740477
step: 10, loss: 0.0003210144641343504
step: 20, loss: 8.284624345833436e-05
step: 30, loss: 5.666921060765162e-05
step: 40, loss: 0.00018798469682224095
step: 50, loss: 0.00035264628240838647
step: 60, loss: 0.0017692585242912173
step: 70, loss: 0.011910998262465
step: 80, loss: 0.00010453650611452758
step: 90, loss: 0.1624298244714737
step: 100, loss: 9.249731374438852e-05
step: 110, loss: 0.000359402853064239
step: 120, loss: 0.00026776420418173075
step: 130, loss: 0.20102393627166748
step: 140, loss: 0.05150889605283737
step: 150, loss: 0.0007277671247720718
step: 160, loss: 0.00018399112741462886
step: 170, loss: 0.0001166264628409408
step: 180, loss: 0.0010640399996191263

```

```

step: 190, loss: 0.0056275962851941586
step: 200, loss: 7.39118258934468e-05
step: 210, loss: 0.0006201340584084392
step: 220, loss: 5.46658193343319e-05
step: 230, loss: 0.0010724675375968218
/usr/local/lib/python3.7/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=0.971815107102593, f1=0.9730337078651685,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000323301472235471
step: 10, loss: 0.00018809171160683036
step: 20, loss: 0.0016561378724873066
step: 30, loss: 0.00024784551351331174
step: 40, loss: 0.0006026406772434711
step: 50, loss: 8.779353811405599e-05
step: 60, loss: 0.018883846700191498
step: 70, loss: 9.225823305314407e-05
step: 80, loss: 9.603537910152227e-05
step: 90, loss: 0.0022109372075647116
step: 100, loss: 8.008706208784133e-05
step: 110, loss: 0.0026093923952430487
step: 120, loss: 0.00025670029572211206
step: 130, loss: 0.029682382941246033
step: 140, loss: 0.00038189050974324346
step: 150, loss: 0.000542407447937876
step: 160, loss: 0.0004498883499763906
step: 170, loss: 0.014008416794240475
step: 180, loss: 0.0013530951691791415
step: 190, loss: 0.00014350384299177676
step: 200, loss: 0.003944468684494495
step: 210, loss: 0.00010442308121128008
step: 220, loss: 6.293720798566937e-05
step: 230, loss: 0.0002497489913366735
/usr/local/lib/python3.7/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=0.9697648376259798, f1=0.9675977653631285,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.2607716750353575e-05
step: 10, loss: 3.255815681768581e-05
step: 20, loss: 3.379483678145334e-05

```

```

step: 30, loss: 0.0001961606030818075
step: 40, loss: 0.02700050175189972
step: 50, loss: 0.00812806561589241
step: 60, loss: 0.0001088818462449126
step: 70, loss: 4.067032932653092e-05
step: 80, loss: 0.0007345355115830898
step: 90, loss: 3.220035432605073e-05
step: 100, loss: 6.636677426286042e-05
step: 110, loss: 9.793305798666552e-05
step: 120, loss: 3.7396763218566775e-05
step: 130, loss: 0.0020787129178643227
step: 140, loss: 4.619377796188928e-05
step: 150, loss: 5.465992580866441e-05
step: 160, loss: 0.0011786342365667224
step: 170, loss: 5.881343167857267e-05
step: 180, loss: 9.499589214101434e-05
step: 190, loss: 0.0019105264218524098
step: 200, loss: 2.6873707611230202e-05
step: 210, loss: 5.6233679060824215e-05
step: 220, loss: 0.0002287783718202263
step: 230, loss: 0.0005048842867836356
/usr/local/lib/python3.7/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=0.9718785151856018, f1=0.971815107102593,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018060224829241633
step: 10, loss: 0.0003901920572388917
step: 20, loss: 0.0002507884637452662
step: 30, loss: 0.00026665054610930383
step: 40, loss: 0.01379481703042984
step: 50, loss: 0.0002225661010015756
step: 60, loss: 0.00010283275332767516
step: 70, loss: 5.9374240663601086e-05
step: 80, loss: 0.0004098328936379403
step: 90, loss: 0.001698615145869553
step: 100, loss: 4.133183028898202e-05
step: 110, loss: 0.002855920698493719
step: 120, loss: 0.0001498542696936056
step: 130, loss: 5.1010098104598e-05
step: 140, loss: 0.00014884081610944122
step: 150, loss: 3.464774999883957e-05
step: 160, loss: 0.0001313384564127773
step: 170, loss: 0.0003192926524206996
step: 180, loss: 0.00011157651169924065

```



```

step: 190, loss: 0.0006113774725235999
step: 200, loss: 0.0004289615317247808
step: 210, loss: 3.76870266336482e-05
step: 220, loss: 6.556395237566903e-05
step: 230, loss: 3.3548865758348256e-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=0.9719416386083053, f1=0.9730337078651685,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.423919694498181e-05
step: 10, loss: 0.00019713552319444716
step: 20, loss: 5.178903302294202e-05
step: 30, loss: 0.04038596898317337
step: 40, loss: 9.756885265232995e-05
step: 50, loss: 0.0002518056717235595
step: 60, loss: 0.0002736650058068335
step: 70, loss: 4.931554940412752e-05
step: 80, loss: 3.8804988435003906e-05
step: 90, loss: 0.00011286399967502803
step: 100, loss: 6.661291263299063e-05
step: 110, loss: 0.05214785039424896
step: 120, loss: 2.5607032512198202e-05
step: 130, loss: 3.630170976975933e-05
step: 140, loss: 0.00018053695384878665
step: 150, loss: 3.248713619541377e-05
step: 160, loss: 0.0004969835863448679
step: 170, loss: 2.518605651857797e-05
step: 180, loss: 5.063816570327617e-05
step: 190, loss: 4.941482620779425e-05
step: 200, loss: 2.9883560273447074e-05
step: 210, loss: 0.0002568190102465451
step: 220, loss: 9.981904440792277e-05
step: 230, loss: 0.0014435243792831898
/usr/local/lib/python3.7/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=0.9708520179372198, f1=0.9718785151856018,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.492681935313158e-05
step: 10, loss: 2.1516918423003517e-05
step: 20, loss: 4.821339825866744e-05

```

```

step: 30, loss: 5.0848713726736605e-05
step: 40, loss: 9.025206963997334e-05
step: 50, loss: 0.001175577170215547
step: 60, loss: 5.417434294940904e-05
step: 70, loss: 8.992914081318304e-05
step: 80, loss: 5.990966747049242e-05
step: 90, loss: 0.0023938759695738554
step: 100, loss: 0.00015829778567422181
step: 110, loss: 3.19134742312599e-05
step: 120, loss: 0.12868383526802063
step: 130, loss: 0.00025752364308573306
step: 140, loss: 0.0012483384925872087
step: 150, loss: 0.005727658048272133
step: 160, loss: 3.559375545592047e-05
step: 170, loss: 3.568223837646656e-05
step: 180, loss: 0.00011589650239329785
step: 190, loss: 0.0009703502291813493
step: 200, loss: 0.0003532047267071903
step: 210, loss: 8.681730832904577e-05
step: 220, loss: 0.00026345718652009964
step: 230, loss: 3.5023105738218874e-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=0.9668141592920354, f1=0.96875, best_f1=0.9698324022346367
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 [00:14, 165.00it/s]
load_f1 = 0.9722530521642618
real_f1 = 0.9680968096809681
0it [00:00, ?it/s]/usr/local/lib/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, 187.78it/s]

```

4.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

```
--lm bert-base-uncased \  
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order:

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

in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

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

step: 0, loss: 0.618700385093689

step: 10, loss: 0.5614398121833801

step: 20, loss: 0.5461257696151733

step: 30, loss: 0.1740122139453888

step: 40, loss: 0.27917981147766113

step: 50, loss: 0.2693474590778351

step: 60, loss: 0.1564817875623703
step: 70, loss: 0.1353410929441452
step: 80, loss: 0.04479261487722397
step: 90, loss: 0.2817336618900299
step: 100, loss: 0.0515357069671154
step: 110, loss: 0.10646023601293564
step: 120, loss: 0.15202854573726654
step: 130, loss: 0.07258746773004532
step: 140, loss: 0.045574430376291275
step: 150, loss: 0.10359257459640503
step: 160, loss: 0.0291945468634367
step: 170, loss: 0.10678312927484512
step: 180, loss: 0.051694776862859726
step: 190, loss: 0.02783849462866783
step: 200, loss: 0.23361320793628693
step: 210, loss: 0.09063762426376343
step: 220, loss: 0.24368062615394592
step: 230, loss: 0.18549492955207825
step: 240, loss: 0.07017293572425842
step: 250, loss: 0.027727706357836723
step: 260, loss: 0.26264750957489014
step: 270, loss: 0.016420602798461914
step: 280, loss: 0.13684388995170593
step: 290, loss: 0.08112350106239319
step: 300, loss: 0.032546937465667725
step: 310, loss: 0.267678439617157
step: 320, loss: 0.13558155298233032
step: 330, loss: 0.020228460431098938
step: 340, loss: 0.13220350444316864
step: 350, loss: 0.07829714566469193
step: 360, loss: 0.14506080746650696
step: 370, loss: 0.06501881033182144
step: 380, loss: 0.02317921817302704
step: 390, loss: 0.25078055262565613
step: 400, loss: 0.29751288890838623
step: 410, loss: 0.06220468506217003
step: 420, loss: 0.05885704234242439
step: 430, loss: 0.16979123651981354
step: 440, loss: 0.023518584668636322
step: 450, loss: 0.006838104221969843
step: 460, loss: 0.00678291404619813
step: 470, loss: 0.10593388229608536
step: 480, loss: 0.11395932734012604
step: 490, loss: 0.16642528772354126
step: 500, loss: 0.11195890605449677
step: 510, loss: 0.07632721960544586
step: 520, loss: 0.016878267750144005
step: 530, loss: 0.05463882163167

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.927007299270073, f1=0.9141824751580849,
best_f1=0.9141824751580849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.105843685567379
step: 10, loss: 0.03504972904920578
step: 20, loss: 0.00968001689761877
step: 30, loss: 0.06627768278121948
step: 40, loss: 0.06546559929847717
step: 50, loss: 0.3748311698436737
step: 60, loss: 0.00870563741773367
step: 70, loss: 0.022903231903910637
step: 80, loss: 0.05150853097438812
step: 90, loss: 0.016123205423355103
step: 100, loss: 0.01879134401679039
step: 110, loss: 0.03767635300755501
step: 120, loss: 0.04693350940942764
step: 130, loss: 0.09401283413171768
step: 140, loss: 0.0315183587372303
step: 150, loss: 0.09859275817871094
step: 160, loss: 0.027878155931830406
step: 170, loss: 0.012350095435976982
step: 180, loss: 0.038571275770664215
step: 190, loss: 0.04786062613129616
step: 200, loss: 0.08066578209400177
step: 210, loss: 0.020698580890893936
step: 220, loss: 0.11304876208305359
step: 230, loss: 0.013743606396019459
step: 240, loss: 0.017931323498487473
step: 250, loss: 0.025230474770069122
step: 260, loss: 0.006585944443941116
step: 270, loss: 0.209150031208992
step: 280, loss: 0.1247105672955513
step: 290, loss: 0.05543551966547966
step: 300, loss: 0.1958329826593399
step: 310, loss: 0.017055686563253403
step: 320, loss: 0.06870938092470169
step: 330, loss: 0.045476630330085754
step: 340, loss: 0.13639122247695923
step: 350, loss: 0.0025660013779997826

```

```

step: 360, loss: 0.236331507563591
step: 370, loss: 0.16181598603725433
step: 380, loss: 0.04681059718132019
step: 390, loss: 0.07663474977016449
step: 400, loss: 0.08320010453462601
step: 410, loss: 0.05973627790808678
step: 420, loss: 0.04744767025113106
step: 430, loss: 0.06133251264691353
step: 440, loss: 0.20334234833717346
step: 450, loss: 0.031025584787130356
step: 460, loss: 0.020523732528090477
step: 470, loss: 0.1984894722700119
step: 480, loss: 0.31093844771385193
step: 490, loss: 0.1089722290635109
step: 500, loss: 0.29653897881507874
step: 510, loss: 0.03721025586128235
step: 520, loss: 0.08724202960729599
step: 530, loss: 0.014702517539262772
/usr/local/lib/python3.7/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=0.9263547938860583, f1=0.9230064161319891,
best_f1=0.9141824751580849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0936652421951294
step: 10, loss: 0.03674882650375366
step: 20, loss: 0.05858056619763374
step: 30, loss: 0.0228231530636549
step: 40, loss: 0.025610504671931267
step: 50, loss: 0.08080226182937622
step: 60, loss: 0.2400204986333847
step: 70, loss: 0.0417356863617897
step: 80, loss: 0.019709495827555656
step: 90, loss: 0.004160976968705654
step: 100, loss: 0.1097959578037262
step: 110, loss: 0.0047613163478672504
step: 120, loss: 0.0114564448595047
step: 130, loss: 0.004399972036480904
step: 140, loss: 0.10754971206188202
step: 150, loss: 0.016930751502513885
step: 160, loss: 0.052357908338308334
step: 170, loss: 0.1343376338481903
step: 180, loss: 0.02999570220708847
step: 190, loss: 0.03672958165407181
step: 200, loss: 0.11093272268772125
step: 210, loss: 0.09244631230831146

```

```

step: 220, loss: 0.018754469230771065
step: 230, loss: 0.11027997732162476
step: 240, loss: 0.006591120269149542
step: 250, loss: 0.0329730287194252
step: 260, loss: 0.04106052592396736
step: 270, loss: 0.05995149910449982
step: 280, loss: 0.07404914498329163
step: 290, loss: 0.002951273927465081
step: 300, loss: 0.01846115104854107
step: 310, loss: 0.005136961583048105
step: 320, loss: 0.00883354339748621
step: 330, loss: 0.019737156108021736
step: 340, loss: 0.008336634375154972
step: 350, loss: 0.007021393161267042
step: 360, loss: 0.017980072647333145
step: 370, loss: 0.02304205857217312
step: 380, loss: 0.09780127555131912
step: 390, loss: 0.007334777619689703
step: 400, loss: 0.03214099258184433
step: 410, loss: 0.0061704968102276325
step: 420, loss: 0.20263640582561493
step: 430, loss: 0.007954007014632225
step: 440, loss: 0.00849489402025938
step: 450, loss: 0.07204796373844147
step: 460, loss: 0.006820004899054766
step: 470, loss: 0.05793992057442665
step: 480, loss: 0.014484894461929798
step: 490, loss: 0.0027940040454268456
step: 500, loss: 0.05275016650557518
step: 510, loss: 0.009221095591783524
step: 520, loss: 0.12965860962867737
step: 530, loss: 0.040603358298540115
/usr/local/lib/python3.7/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=0.9256351039260969, f1=0.9220839096357768,
best_f1=0.9141824751580849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010142640210688114
step: 10, loss: 0.007704328279942274
step: 20, loss: 0.002269577933475375
step: 30, loss: 0.028245646506547928
step: 40, loss: 0.05811984837055206
step: 50, loss: 0.05762828141450882
step: 60, loss: 0.0024816447403281927
step: 70, loss: 0.0014498134842142463

```


step: 80, loss: 0.01015451643615961
step: 90, loss: 0.086692214012146
step: 100, loss: 0.007638911250978708
step: 110, loss: 0.0245196595788002
step: 120, loss: 0.0018123133340850472
step: 130, loss: 0.02402574010193348
step: 140, loss: 0.0011524477740749717
step: 150, loss: 0.007125609088689089
step: 160, loss: 0.08522679656744003
step: 170, loss: 0.03004498966038227
step: 180, loss: 0.002985801314935088
step: 190, loss: 0.04876473918557167
step: 200, loss: 0.020334597676992416
step: 210, loss: 0.03751075267791748
step: 220, loss: 0.011197131127119064
step: 230, loss: 0.01567365974187851
step: 240, loss: 0.00287642446346581
step: 250, loss: 0.01654943823814392
step: 260, loss: 0.024311719462275505
step: 270, loss: 0.057954974472522736
step: 280, loss: 0.035346124321222305
step: 290, loss: 0.015125189907848835
step: 300, loss: 0.002296044258400798
step: 310, loss: 0.0221333559602499
step: 320, loss: 0.1582615226507187
step: 330, loss: 0.010866372846066952
step: 340, loss: 0.009452376514673233
step: 350, loss: 0.02068687416613102
step: 360, loss: 0.01961854100227356
step: 370, loss: 0.0012927247444167733
step: 380, loss: 0.009336923249065876
step: 390, loss: 0.02729814499616623
step: 400, loss: 0.003775420831516385
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 410, loss: 0.010906262323260307
step: 420, loss: 0.0016845718491822481
step: 430, loss: 0.03173334151506424
step: 440, loss: 0.0007873232825659215
step: 450, loss: 0.011926177889108658
step: 460, loss: 0.014146527275443077
step: 470, loss: 0.05653833970427513
step: 480, loss: 0.05092550441622734
step: 490, loss: 0.026542935520410538
step: 500, loss: 0.023336263373494148
step: 510, loss: 0.040941882878541946
step: 520, loss: 0.02035684883594513
step: 530, loss: 0.04999515041708946
/usr/local/lib/python3.7/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=0.9142857142857143, f1=0.9188687992582291,
best_f1=0.9141824751580849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06417515128850937
step: 10, loss: 0.04379195347428322
step: 20, loss: 0.020105566829442978
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.00835246779024601
step: 40, loss: 0.16034923493862152
step: 50, loss: 0.006883797235786915
step: 60, loss: 0.05268925055861473
step: 70, loss: 0.06747901439666748
step: 80, loss: 0.04125412553548813
step: 90, loss: 0.14058010280132294
step: 100, loss: 0.011750360950827599
step: 110, loss: 0.02921309694647789
step: 120, loss: 0.011807815171778202
step: 130, loss: 0.00037482805782929063
step: 140, loss: 0.03533203527331352
step: 150, loss: 0.005545677151530981
step: 160, loss: 0.0006754976930096745
step: 170, loss: 0.03209812194108963
step: 180, loss: 0.0009107619407586753
step: 190, loss: 0.0032786971423774958
step: 200, loss: 0.0024023978039622307
step: 210, loss: 0.017179924994707108
step: 220, loss: 0.002941797487437725
step: 230, loss: 0.03273797407746315
step: 240, loss: 0.023622894659638405
step: 250, loss: 0.0013098445488139987
step: 260, loss: 0.006166487466543913
step: 270, loss: 0.01102522388100624
step: 280, loss: 0.011682097800076008
step: 290, loss: 0.1327064335346222
step: 300, loss: 0.013782848604023457
step: 310, loss: 0.016975848004221916
step: 320, loss: 0.143433079123497
step: 330, loss: 0.01484184805303812
step: 340, loss: 0.02684483677148819
step: 350, loss: 0.011110990308225155
step: 360, loss: 0.00947115384042263
step: 370, loss: 0.05410310998558998
step: 380, loss: 0.0015862329164519906
step: 390, loss: 0.0015593208372592926
```

```

step: 400, loss: 0.010566622018814087
step: 410, loss: 0.004476530011743307
step: 420, loss: 0.0052252947352826595
step: 430, loss: 0.01692512258887291
step: 440, loss: 0.007173636928200722
step: 450, loss: 0.006948635447770357
step: 460, loss: 0.0014695122372359037
step: 470, loss: 0.006340933032333851
step: 480, loss: 0.019637389108538628
step: 490, loss: 0.0061553162522614
step: 500, loss: 0.053357481956481934
step: 510, loss: 0.11541422456502914
step: 520, loss: 0.0008647013455629349
step: 530, loss: 0.0010585386771708727
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.930776426566885, f1=0.9253592953175707,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003735172562301159
step: 10, loss: 0.0010196908842772245
step: 20, loss: 0.16913604736328125
step: 30, loss: 0.007643583230674267
step: 40, loss: 0.02080732397735119
step: 50, loss: 0.02495582401752472
step: 60, loss: 0.00046426369226537645
step: 70, loss: 0.0023700513411313295
step: 80, loss: 0.0008251843973994255
step: 90, loss: 0.0002984409511554986
step: 100, loss: 0.012269491329789162
step: 110, loss: 0.00219930917955935
step: 120, loss: 0.0007457408355548978
step: 130, loss: 0.018848421052098274
step: 140, loss: 0.014537995681166649
step: 150, loss: 0.02660335600376129
step: 160, loss: 0.0015422486467286944
step: 170, loss: 0.024543972685933113
step: 180, loss: 0.0027553788386285305
step: 190, loss: 0.0006858441629447043
step: 200, loss: 0.0007209375617094338
step: 210, loss: 0.005547880660742521

```

```

step: 220, loss: 0.008597269654273987
step: 230, loss: 0.007820634171366692
step: 240, loss: 0.11953180283308029
step: 250, loss: 0.03896770626306534
step: 260, loss: 0.015596749261021614
step: 270, loss: 0.12049908936023712
step: 280, loss: 0.028954030945897102
step: 290, loss: 0.0012707783607766032
step: 300, loss: 0.06456232070922852
step: 310, loss: 0.021747222170233727
step: 320, loss: 0.0020319742616266012
step: 330, loss: 0.0029097208753228188
step: 340, loss: 0.03078596107661724
step: 350, loss: 0.026252472773194313
step: 360, loss: 0.013028619810938835
step: 370, loss: 0.08744469285011292
step: 380, loss: 0.002337908372282982
step: 390, loss: 0.034590963274240494
step: 400, loss: 0.00955025665462017
step: 410, loss: 0.003364310832694173
step: 420, loss: 0.13988789916038513
step: 430, loss: 0.014989808201789856
step: 440, loss: 0.0006906575872562826
step: 450, loss: 0.00034807337215170264
step: 460, loss: 0.0010675222147256136
step: 470, loss: 0.012440245598554611
step: 480, loss: 0.005248433910310268
step: 490, loss: 0.005360874813050032
step: 500, loss: 0.000555452483240515
step: 510, loss: 0.005468061193823814
step: 520, loss: 0.014390927739441395
step: 530, loss: 0.0049340995028615
/usr/local/lib/python3.7/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=0.9290023201856149, f1=0.927816091954023,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002802594972308725
step: 10, loss: 0.09492655098438263
step: 20, loss: 0.010716949589550495
step: 30, loss: 0.0015873115044087172
step: 40, loss: 0.18455898761749268
step: 50, loss: 0.0057993740774691105
step: 60, loss: 0.029523732140660286
step: 70, loss: 0.06655406951904297

```

step: 80, loss: 0.004533145576715469
step: 90, loss: 0.0007401246693916619
step: 100, loss: 0.0008797469781711698
step: 110, loss: 0.0007768162176944315
step: 120, loss: 0.00365068088285625
step: 130, loss: 0.015690382570028305
step: 140, loss: 0.00020414101891219616
step: 150, loss: 0.001236048061400652
step: 160, loss: 0.0002658895682543516
step: 170, loss: 0.0035827532410621643
step: 180, loss: 0.0036359878722578287
step: 190, loss: 0.01149247121065855
step: 200, loss: 0.01341770775616169
step: 210, loss: 0.008245772682130337
step: 220, loss: 0.003866571234539151
step: 230, loss: 0.009002666920423508
step: 240, loss: 0.0005817727069370449
step: 250, loss: 0.002173110144212842
step: 260, loss: 0.0005738483159802854
step: 270, loss: 0.015753336250782013
step: 280, loss: 0.0011317674070596695
step: 290, loss: 0.00039235755684785545
step: 300, loss: 0.003969650249928236
step: 310, loss: 0.00027774323825724423
step: 320, loss: 0.0047123911790549755
step: 330, loss: 0.0042886147275567055
step: 340, loss: 0.07587121427059174
step: 350, loss: 0.001622502226382494
step: 360, loss: 0.0014760529156774282
step: 370, loss: 0.0014826323604211211
step: 380, loss: 0.00046336735249496996
step: 390, loss: 0.0029822001233696938
step: 400, loss: 0.013172419741749763
step: 410, loss: 0.005488757509738207
step: 420, loss: 0.0025548646226525307
step: 430, loss: 0.005247634369879961
step: 440, loss: 0.00036987633211538196
step: 450, loss: 0.09929441660642624
step: 460, loss: 0.004164435435086489
step: 470, loss: 0.004095442593097687
step: 480, loss: 0.16965164244174957
step: 490, loss: 0.003250280860811472
step: 500, loss: 0.0335354246199131
step: 510, loss: 0.0009763186681084335
step: 520, loss: 0.0597991868853569
step: 530, loss: 0.11348079144954681
/usr/local/lib/python3.7/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=0.9238625812441968, f1=0.9225806451612903,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07125703990459442
step: 10, loss: 0.007582966238260269
step: 20, loss: 0.0004876853490713984
step: 30, loss: 0.0073622711934149265
step: 40, loss: 0.00011349366832291707
step: 50, loss: 0.0005133224185556173
step: 60, loss: 0.0005149643984623253
step: 70, loss: 0.002630525967106223
step: 80, loss: 0.01021440140902996
step: 90, loss: 0.0009403491858392954
step: 100, loss: 0.00039798280340619385
step: 110, loss: 0.0019207693403586745
step: 120, loss: 0.010140829719603062
step: 130, loss: 0.0026002563536167145
step: 140, loss: 0.013227426446974277
step: 150, loss: 0.00023678317666053772
step: 160, loss: 0.007666042540222406
step: 170, loss: 0.004894290119409561
step: 180, loss: 0.0009700586088001728
step: 190, loss: 0.0011906888103112578
step: 200, loss: 0.00045716940076090395
step: 210, loss: 0.0002975817769765854
step: 220, loss: 0.005496541038155556
step: 230, loss: 0.001127389376051724
step: 240, loss: 0.00031834241235628724
step: 250, loss: 0.00013711597421206534
step: 260, loss: 0.00015799920947756618
step: 270, loss: 0.0005136698018759489
step: 280, loss: 0.01712401956319809
step: 290, loss: 7.355395064223558e-05
step: 300, loss: 0.0002661336911842227
step: 310, loss: 0.07009257376194
step: 320, loss: 0.014961345121264458
step: 330, loss: 0.003535522846505046
step: 340, loss: 0.00011786259710788727
step: 350, loss: 0.0004319175495766103
step: 360, loss: 0.0012810623738914728
step: 370, loss: 0.0007839316967874765
step: 380, loss: 0.000609010923653841
step: 390, loss: 0.12129493057727814
step: 400, loss: 0.005817974917590618
step: 410, loss: 0.00022749151685275137
```

```
step: 420, loss: 0.00398798706009984
step: 430, loss: 0.031820327043533325
step: 440, loss: 0.004171355161815882
step: 450, loss: 0.00020485284039750695
step: 460, loss: 0.035399239510297775
step: 470, loss: 0.0008490199106745422
step: 480, loss: 0.00044512402382679284
step: 490, loss: 0.02075447514653206
step: 500, loss: 0.0007116476772353053
step: 510, loss: 0.00040962608181871474
step: 520, loss: 0.0017349130939692259
step: 530, loss: 0.0006282085087150335
/usr/local/lib/python3.7/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=0.9258741258741259, f1=0.9249304911955515,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020623779855668545
step: 10, loss: 0.002390738343819976
step: 20, loss: 0.019320979714393616
step: 30, loss: 0.002144654281437397
step: 40, loss: 0.004505406133830547
step: 50, loss: 0.0009055180125869811
step: 60, loss: 0.0073872958309948444
step: 70, loss: 0.0010428267996758223
step: 80, loss: 0.030997779220342636
step: 90, loss: 0.00022143212845548987
step: 100, loss: 0.000130363623611629
step: 110, loss: 0.0016787817003205419
step: 120, loss: 0.006248593796044588
step: 130, loss: 0.048824869096279144
step: 140, loss: 0.0003224418614991009
step: 150, loss: 0.00041125292773358524
step: 160, loss: 0.0008197513525374234
step: 170, loss: 0.021807575598359108
step: 180, loss: 0.0006784126744605601
step: 190, loss: 0.10297657549381256
step: 200, loss: 0.0015653896844014525
step: 210, loss: 0.0004166875733062625
step: 220, loss: 0.03534988313913345
step: 230, loss: 0.0007889136904850602
step: 240, loss: 0.0005243090563453734
step: 250, loss: 0.001061331364326179
step: 260, loss: 0.0023274088744074106
step: 270, loss: 6.422296428354457e-05
```

```

step: 280, loss: 0.009050650522112846
step: 290, loss: 0.02842411957681179
step: 300, loss: 0.0002473421045579016
step: 310, loss: 0.0029707755893468857
step: 320, loss: 0.0009265571134164929
step: 330, loss: 0.00032238528365269303
step: 340, loss: 0.001832769950851798
step: 350, loss: 0.09585762768983841
step: 360, loss: 6.204860255820677e-05
step: 370, loss: 0.00652709323912859
step: 380, loss: 0.002990598790347576
step: 390, loss: 0.0002997700939886272
step: 400, loss: 0.0016793792601674795
step: 410, loss: 0.004501967690885067
step: 420, loss: 0.06912454217672348
step: 430, loss: 0.0025713187642395496
step: 440, loss: 0.004800120368599892
step: 450, loss: 0.0037696848157793283
step: 460, loss: 0.001186639186926186
step: 470, loss: 0.00014204013859853148
step: 480, loss: 0.002808500314131379
step: 490, loss: 0.0252737607806921
step: 500, loss: 0.0028086828533560038
step: 510, loss: 0.0033756857737898827
step: 520, loss: 0.0015740649541839957
step: 530, loss: 0.0026536916848272085
/usr/local/lib/python3.7/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=0.9307400379506643, f1=0.9230046948356808,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032645168248564005
step: 10, loss: 0.002136706607416272
step: 20, loss: 6.177269096951932e-05
step: 30, loss: 0.00548803573474288
step: 40, loss: 0.0004929768037982285
step: 50, loss: 0.0003609923296608031
step: 60, loss: 0.00010120764636667445
step: 70, loss: 0.00023605884052813053
step: 80, loss: 4.588976298691705e-05
step: 90, loss: 0.00010273310908814892
step: 100, loss: 0.0007044916274026036
step: 110, loss: 0.0006091844406910241
step: 120, loss: 0.000968819425906986
step: 130, loss: 0.0004164606798440218

```



```

step: 140, loss: 0.0007215681253001094
step: 150, loss: 0.00029459656798280776
step: 160, loss: 0.00015664330567233264
step: 170, loss: 0.002428567036986351
step: 180, loss: 0.00043033668771386147
step: 190, loss: 0.0006938343867659569
step: 200, loss: 0.000554974190890789
step: 210, loss: 0.0016252173809334636
step: 220, loss: 0.0037429414223879576
step: 230, loss: 0.0002958819968625903
step: 240, loss: 0.06252893060445786
step: 250, loss: 0.0014800394419580698
step: 260, loss: 0.004055251833051443
step: 270, loss: 0.0020665228366851807
step: 280, loss: 0.0001953902537934482
step: 290, loss: 0.00025095255114138126
step: 300, loss: 5.574140232056379e-05
step: 310, loss: 0.0003412330988794565
step: 320, loss: 0.0018183219945058227
step: 330, loss: 0.0031021242029964924
step: 340, loss: 0.009594817645847797
step: 350, loss: 0.00012424025044310838
step: 360, loss: 0.00014111165364738554
step: 370, loss: 0.01669926382601261
step: 380, loss: 0.002799797337502241
step: 390, loss: 0.0003164478694088757
step: 400, loss: 0.0024382807314395905
step: 410, loss: 0.0007727727061137557
step: 420, loss: 0.00041342712938785553
step: 430, loss: 3.3653366699581966e-05
step: 440, loss: 0.003951305989176035
step: 450, loss: 9.467772179050371e-05
step: 460, loss: 0.000847774965222925
step: 470, loss: 0.0015897291013970971
step: 480, loss: 0.00012633255391847342
step: 490, loss: 0.028184570372104645
step: 500, loss: 0.025975322350859642
step: 510, loss: 9.499111911281943e-05
step: 520, loss: 0.07335054874420166
step: 530, loss: 0.0004203228163532913
/usr/local/lib/python3.7/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=0.9244655581947744, f1=0.914339801230478,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.11630771309137344
step: 10, loss: 0.05601054057478905
step: 20, loss: 0.0008260325994342566
step: 30, loss: 0.0010878773173317313
step: 40, loss: 9.935154230333865e-05
step: 50, loss: 0.000886164722032845
step: 60, loss: 0.023564651608467102
step: 70, loss: 7.450301200151443e-05
step: 80, loss: 0.0026079858653247356
step: 90, loss: 0.0003323773853480816
step: 100, loss: 0.0009710274753160775
step: 110, loss: 0.0002206778444815427
step: 120, loss: 0.00027887721080332994
step: 130, loss: 9.617396426619962e-05
step: 140, loss: 0.0004873817088082433
step: 150, loss: 0.00026014912873506546
step: 160, loss: 0.015328721143305302
step: 170, loss: 0.0006138734752312303
step: 180, loss: 0.0014625893672928214
step: 190, loss: 0.0015228866832330823
step: 200, loss: 0.017183193936944008
step: 210, loss: 0.00032076635397970676
step: 220, loss: 0.003800460370257497
step: 230, loss: 0.00013458479952532798
step: 240, loss: 0.0005346059915609658
step: 250, loss: 8.101824641926214e-05
step: 260, loss: 0.0006165506201796234
step: 270, loss: 0.0016862592892721295
step: 280, loss: 0.004093383904546499
step: 290, loss: 0.011841334402561188
step: 300, loss: 0.00012525530473794788
step: 310, loss: 0.007991451770067215
step: 320, loss: 0.00020939047681167722
step: 330, loss: 0.0005796545883640647
step: 340, loss: 0.15000832080841064
step: 350, loss: 0.000711932429112494
step: 360, loss: 0.00020016092457808554
step: 370, loss: 0.05838671699166298
step: 380, loss: 0.00022003961203154176
step: 390, loss: 4.6238594222813845e-05
step: 400, loss: 0.0005559687851928174
step: 410, loss: 0.0012924842303618789
step: 420, loss: 9.714106272440404e-05
step: 430, loss: 0.002527889097109437
step: 440, loss: 0.0004947755369357765
step: 450, loss: 6.295581988524646e-05
step: 460, loss: 0.17180253565311432
step: 470, loss: 0.0008858618675731122

```

step: 480, loss: 0.0014654018450528383
step: 490, loss: 0.0001652810169616714
step: 500, loss: 0.00016387640789616853
step: 510, loss: 0.28440025448799133
step: 520, loss: 0.00010818035661941394
step: 530, loss: 0.0050632543861866
/usr/local/lib/python3.7/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=0.929472209248015, f1=0.9229349330872173,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003435168764553964
step: 10, loss: 0.00018072070088237524
step: 20, loss: 0.0002760059724096209
step: 30, loss: 0.001606079749763012
step: 40, loss: 0.0008552793296985328
step: 50, loss: 0.021862080320715904
step: 60, loss: 0.000677716510836035
step: 70, loss: 0.21611981093883514
step: 80, loss: 8.565410098526627e-05
step: 90, loss: 0.00014038392691873014
step: 100, loss: 0.010507945902645588
step: 110, loss: 0.012512262910604477
step: 120, loss: 0.000254020094871521
step: 130, loss: 0.0004073190502822399
step: 140, loss: 0.0015007617184892297
step: 150, loss: 0.00013697502436116338
step: 160, loss: 4.194087159703486e-05
step: 170, loss: 6.879209104226902e-05
step: 180, loss: 0.0001456805766792968
step: 190, loss: 0.0015242714434862137
step: 200, loss: 0.00130283844191581
step: 210, loss: 0.0005671142716892064
step: 220, loss: 0.0007263334118761122
step: 230, loss: 9.164701623376459e-05
step: 240, loss: 0.016218386590480804
step: 250, loss: 0.002360948594287038
step: 260, loss: 0.00014178194396663457
step: 270, loss: 0.00033189228270202875
step: 280, loss: 0.0011476089712232351
step: 290, loss: 0.01573793590068817
step: 300, loss: 0.0033711157739162445
step: 310, loss: 0.0004121558158658445
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 320, loss: 0.000325383065501228

```

```

step: 330, loss: 0.01002555899322033
step: 340, loss: 0.0002459487586747855
step: 350, loss: 0.00041692418744787574
step: 360, loss: 0.0638628751039505
step: 370, loss: 0.007102236617356539
step: 380, loss: 0.0008866747375577688
step: 390, loss: 0.00038648193003609776
step: 400, loss: 0.0003663812531158328
step: 410, loss: 0.02494376339018345
step: 420, loss: 0.009154146537184715
step: 430, loss: 0.0076905726455152035
step: 440, loss: 0.014216159470379353
step: 450, loss: 0.0005496313679032028
step: 460, loss: 8.85050103534013e-05
step: 470, loss: 0.0039972299709916115
step: 480, loss: 0.0036317480262368917
step: 490, loss: 0.0014101001434028149
step: 500, loss: 0.00014664135233033448
step: 510, loss: 0.00024537567514926195
step: 520, loss: 0.00010009911056840792
step: 530, loss: 0.0011251146206632257
/usr/local/lib/python3.7/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=0.9306466729147141, f1=0.9240801117838845,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004371055401861668
step: 10, loss: 0.00041161334956996143
step: 20, loss: 0.001330193947069347
step: 30, loss: 0.0004885743255726993
step: 40, loss: 0.002182903466746211
step: 50, loss: 0.03248502314090729
step: 60, loss: 0.00030678973416797817
step: 70, loss: 6.861241126898676e-05
step: 80, loss: 0.0002257432643091306
step: 90, loss: 0.014522590674459934
step: 100, loss: 0.0037383525632321835
step: 110, loss: 0.001264961902052164
step: 120, loss: 0.00019653343770187348
step: 130, loss: 0.00010615348583087325
step: 140, loss: 0.00106245547067374
step: 150, loss: 0.0006440674769692123
step: 160, loss: 0.002091581467539072
step: 170, loss: 0.0012490007793530822
step: 180, loss: 0.00018867706239689142

```

```
step: 190, loss: 0.00011771250865422189
step: 200, loss: 5.199189763516188e-05
step: 210, loss: 0.0001281442673644051
step: 220, loss: 5.5439155403291807e-05
step: 230, loss: 0.0011442701797932386
step: 240, loss: 0.0019161589443683624
step: 250, loss: 6.866119656478986e-05
step: 260, loss: 0.000104957063740585
step: 270, loss: 4.1591058106860146e-05
step: 280, loss: 3.408532575122081e-05
step: 290, loss: 0.0007202302222140133
step: 300, loss: 0.0004779379814863205
step: 310, loss: 0.0017888476140797138
step: 320, loss: 0.003923504147678614
step: 330, loss: 0.003070783568546176
step: 340, loss: 0.0005057572852820158
step: 350, loss: 0.00010130214650416747
step: 360, loss: 4.386641739984043e-05
step: 370, loss: 0.0020759638864547014
step: 380, loss: 8.295763109344989e-05
step: 390, loss: 0.00031214652699418366
step: 400, loss: 0.00010028838005382568
step: 410, loss: 0.0009998814202845097
step: 420, loss: 0.0006982627674005926
step: 430, loss: 0.03902564197778702
step: 440, loss: 0.0008175880066119134
step: 450, loss: 9.791112825041637e-05
step: 460, loss: 0.002411814406514168
step: 470, loss: 6.625625974265859e-05
step: 480, loss: 0.0002560386201366782
step: 490, loss: 0.00018444065062794834
step: 500, loss: 0.00017434160690754652
step: 510, loss: 0.0021615817677229643
step: 520, loss: 0.00287647545337677
step: 530, loss: 0.0001757394493324682
/usr/local/lib/python3.7/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=0.9301675977653632, f1=0.9301895515487749,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009619179181754589
step: 10, loss: 0.0007544214022345841
step: 20, loss: 0.00016781855083536357
step: 30, loss: 4.63544565718621e-05
step: 40, loss: 0.0005609671352431178
```

step: 50, loss: 0.009692889638245106
step: 60, loss: 3.525818101479672e-05
step: 70, loss: 0.006105988752096891
step: 80, loss: 0.00010165087587665766
step: 90, loss: 0.0002483779680915177
step: 100, loss: 0.0001946914562722668
step: 110, loss: 7.942268712213263e-05
step: 120, loss: 0.00019914157746825367
step: 130, loss: 0.0772453173995018
step: 140, loss: 0.0016889465041458607
step: 150, loss: 0.00013562718231696635
step: 160, loss: 7.440039189532399e-05
step: 170, loss: 0.0004277440602891147
step: 180, loss: 4.293033271096647e-05
step: 190, loss: 6.732853216817603e-05
step: 200, loss: 0.00020373753795865923
step: 210, loss: 5.9676058299373835e-05
step: 220, loss: 6.449734064517543e-05
step: 230, loss: 0.014926704578101635
step: 240, loss: 0.0001751439122017473
step: 250, loss: 0.00010579806985333562
step: 260, loss: 0.0004467099788598716
step: 270, loss: 0.00021437476971186697
step: 280, loss: 4.5586763008031994e-05
step: 290, loss: 0.00350035703741014
step: 300, loss: 0.00183330406434834
step: 310, loss: 0.0009799012914299965
step: 320, loss: 0.00022948755940888077
step: 330, loss: 0.00021687991102226079
step: 340, loss: 0.0006763795972801745
step: 350, loss: 0.0017931752372533083
step: 360, loss: 0.14666420221328735
step: 370, loss: 0.00025708068278618157
step: 380, loss: 0.0002716721792239696
step: 390, loss: 0.009840010665357113
step: 400, loss: 0.00017997990653384477
step: 410, loss: 3.4073796996381134e-05
step: 420, loss: 0.00017696726717986166
step: 430, loss: 0.007515540346503258
step: 440, loss: 0.00022815327974967659
step: 450, loss: 0.00016943579248618335
step: 460, loss: 0.000771188351791352
step: 470, loss: 7.214558718260378e-05
step: 480, loss: 0.00033448138856329024
step: 490, loss: 3.9065540477167815e-05
step: 500, loss: 6.14953605690971e-05
step: 510, loss: 0.002636552322655916
step: 520, loss: 0.0003257080679759383

```
step: 530, loss: 0.000228555261855945
/usr/local/lib/python3.7/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=0.927468413664015, f1=0.9265116279069768,
best_f1=0.9253592953175707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.6260924136731774e-05
step: 10, loss: 0.00019741483265534043
step: 20, loss: 0.0003762042324524373
step: 30, loss: 0.00021704503160435706
step: 40, loss: 0.05615336820483208
step: 50, loss: 0.0022023331839591265
step: 60, loss: 6.575667794095352e-05
step: 70, loss: 0.0006662971572950482
step: 80, loss: 0.00013376543938647956
step: 90, loss: 0.00026163176517002285
step: 100, loss: 0.00025046372320502996
step: 110, loss: 0.0006276360945776105
step: 120, loss: 0.00011259944585617632
step: 130, loss: 0.04782569408416748
step: 140, loss: 0.00011061136319767684
step: 150, loss: 0.007948391139507294
step: 160, loss: 0.0007050891872495413
step: 170, loss: 0.00011001870007021353
step: 180, loss: 0.00019706397142726928
step: 190, loss: 0.00011676608846755698
step: 200, loss: 0.0012191572459414601
step: 210, loss: 0.0006739292875863612
step: 220, loss: 0.10052353888750076
step: 230, loss: 0.0009910025401040912
step: 240, loss: 0.0017858271021395922
step: 250, loss: 0.0005246654036454856
step: 260, loss: 0.0008687354275025427
step: 270, loss: 4.091996015631594e-05
step: 280, loss: 0.0006496140849776566
step: 290, loss: 4.159032687311992e-05
step: 300, loss: 2.8460357498261146e-05
step: 310, loss: 0.0005345107638277113
step: 320, loss: 0.001380372792482376
step: 330, loss: 7.019913755357265e-05
step: 340, loss: 7.79524416429922e-05
step: 350, loss: 0.00024719006614759564
step: 360, loss: 0.001263865502551198
step: 370, loss: 0.00023267013602890074
step: 380, loss: 0.00030892505310475826
```

```

step: 390, loss: 0.1802985519170761
step: 400, loss: 0.00011087657912867144
step: 410, loss: 0.0005725956871174276
step: 420, loss: 0.00014053111954126507
step: 430, loss: 0.00028704648138955235
step: 440, loss: 0.05134430527687073
step: 450, loss: 4.161951073911041e-05
step: 460, loss: 3.445373658905737e-05
step: 470, loss: 0.006740489974617958
step: 480, loss: 5.427592986961827e-05
step: 490, loss: 3.235874100937508e-05
step: 500, loss: 9.276404307456687e-05
step: 510, loss: 0.00019314276869408786
step: 520, loss: 9.399289410794154e-05
step: 530, loss: 8.34885795484297e-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=0.9285042333019755, f1=0.927468413664015,
best_f1=0.9253592953175707
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 [00:29, 193.78it/s]  
load_f1 = 0.9246607393542348  
real_f1 = 0.9250116441546343  
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-  
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
4398it [00:22, 192.78it/s]
```