

CMedium_30_3_5_bert

August 13, 2022

0.1 Install packages

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 23.28 MiB/s, done.
Resolving deltas: 100% (142/142), done.
Checking out files: 100% (128/128), 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 2.3 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 5.6 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 52.6 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 76.1 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 50.3 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 8.8 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 21.52 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
    | 10.4 MB 4.4 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 69.9 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
    | 753.4 MB 7.0 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
    | 56 kB 5.5 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
    | 884 kB 65.2 MB/s
Collecting jsonlines==1.2.0

```

```

    Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)
Collecting nltk==3.4.5
    Downloading nltk-3.4.5.zip (1.5 MB)
      |                               | 1.5 MB 54.6 MB/s
Requirement already satisfied: scipy>=0.18.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.7.3)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Collecting blis<0.5.0,>=0.4.0
    Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
      |                               | 3.7 MB 52.7 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)
Collecting plac<1.2.0,>=0.9.6
    Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting thinc<7.4.0,>=7.3.0
    Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
      |                               | 2.2 MB 50.5 MB/s
Collecting srsly<1.1.0,>=0.1.0
    Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
      |                               | 184 kB 73.4 MB/s
Collecting catalogue<1.1.0,>=0.0.7
    Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
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: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting tokenizers==0.8.1.rc2
    Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)

```

```

| 3.0 MB 42.4 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    | 880 kB 56.1 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: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!<1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
    Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449923
    sha256=aa652793ccf87d8489567a17a7f4c8d651e91d7c941257851a53d9a749efee6d
    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=1e93cc126502dbc39af619df66254087fce2af90a7c9f03c6e184aee70f0a7a1
Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
Attempting uninstall: numpy
  Found existing installation: numpy 1.21.6
  Uninstalling numpy-1.21.6:
    Successfully uninstalled numpy-1.21.6
Attempting uninstall: tqdm
  Found existing installation: tqdm 4.64.0
  Uninstalling tqdm-4.64.0:
    Successfully uninstalled tqdm-4.64.0
Attempting uninstall: srsly
  Found existing installation: srsly 2.4.4
  Uninstalling srsly-2.4.4:
    Successfully uninstalled srsly-2.4.4
Attempting uninstall: regex
  Found existing installation: regex 2022.6.2
  Uninstalling regex-2022.6.2:
    Successfully uninstalled regex-2022.6.2
Attempting uninstall: blis
  Found existing installation: blis 0.7.8
  Uninstalling blis-0.7.8:
    Successfully uninstalled blis-0.7.8
Attempting uninstall: thinc
  Found existing installation: thinc 8.1.0
  Uninstalling thinc-8.1.0:
    Successfully uninstalled thinc-8.1.0
Attempting uninstall: catalogue
  Found existing installation: catalogue 2.0.8
  Uninstalling catalogue-2.0.8:
    Successfully uninstalled catalogue-2.0.8
Attempting uninstall: torch
  Found existing installation: torch 1.12.0+cu113
  Uninstalling torch-1.12.0+cu113:
    Successfully uninstalled torch-1.12.0+cu113
Attempting uninstall: spacy
  Found existing installation: spacy 3.4.1
  Uninstalling spacy-3.4.1:
    Successfully uninstalled spacy-3.4.1
Attempting uninstall: nltk
  Found existing installation: nltk 3.7
  Uninstalling nltk-3.7:

```

```

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

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

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

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

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

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

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

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

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

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

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

```

```
[1]: True
```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10117, done.
remote: Counting objects: 100% (233/233), done.
remote: Compressing objects: 100% (165/165), done.
remote: Total 10117 (delta 119), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.97 MiB | 22.02 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-pjuwhrsm
Created temporary directory: /tmp/pip-req-tracker-pbilpejy
Initialized build tracking at /tmp/pip-req-tracker-pbilpejy
Created build tracker: /tmp/pip-req-tracker-pbilpejy
Entered build tracker: /tmp/pip-req-tracker-pbilpejy
```

```

Created temporary directory: /tmp/pip-install-62g4y6ax
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-31e_efpy
  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-
pbilpejy'
  Running setup.py (path:/tmp/pip-req-build-31e_efpy/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-8t246188
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-8t246188/apex.egg-info
  writing /tmp/pip-pip-egg-info-8t246188/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-8t246188/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-8t246188/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-8t246188/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-8t246188/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-31e_efpy 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-pbilpejy'
Created temporary directory: /tmp/pip-unpack-d8nknlve
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-sctiqzok
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-sctiqzok
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''''/tmp/pip-req-build-31e_efpy/setup.py'''';

```



```
__file__='''/tmp/pip-req-build-31e_efpy/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-sctiqzok
```

```
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-sctiqzok/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=af39b63dc61d0d0f846a74e3b34fc96c6c6102322ce44c199db8d50e9d9edfd6
Stored in directory: /tmp/pip-ephem-wheel-cache-
pjuwhrsm/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-pbi1pejy'
/content/ditto

```

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                               | 125 kB 4.9 MB/s
Collecting awscli
  Downloading awscli-1.25.49-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 45.5 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 65.9 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting boto3==1.27.49
  Downloading boto3-1.27.49-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 42.7 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 8.8 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
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

```

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

```

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

```

```

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

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

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

```
[5]: !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 | 27.98 MiB/s, done.
Resolving deltas: 100% (635/635), done.
Checking out files: 100% (1304/1304), done.

```

0.4 remove e move

```
[6]: !ls
```

```

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

```

```
[7]: !rm configs.json
!mv -n /content/ditto/IC/datasEsErros/CMedium_30_3_5/configs.json /content/
↪ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

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

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

Downloading: 100% 433/433 [00:00<00:00, 344kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 351kB/s]

Downloading: 100% 440M/440M [00:11<00:00, 39.4MB/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: 0.5635988116264343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.34782608695652173, f1=0.30769230769230765,
best_f1=0.30769230769230765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5061610341072083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.3529411764705882, f1=0.3157894736842105,
best_f1=0.3157894736842105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5584540963172913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.39285714285714285, 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: 0.28716859221458435
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5128205128205129, 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: 0.24942944943904877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6250000000000001, f1=0.5517241379310344,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29641950130462646

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.689655172413793, 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: 0.04376140609383583
/usr/local/lib/python3.7/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=0.689655172413793, 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: 0.01952725276350975
/usr/local/lib/python3.7/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=0.6666666666666666, 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: 0.002095213858410716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7200000000000001, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04278089478611946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```


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=0.7407407407407408, f1=0.5714285714285714,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005829306319355965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.7692307692307692, f1=0.6206896551724138,
best_f1=0.6206896551724138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008200390264391899
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.5925925925925927,
best_f1=0.6206896551724138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005688575562089682
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6153846153846153,
best_f1=0.6206896551724138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005621185526251793
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6153846153846153,
best_f1=0.6206896551724138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
```

```

step: 0, loss: 0.004622721578925848
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6153846153846153,
best_f1=0.6206896551724138
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 65989.22it/s]
/usr/local/lib/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 = 0.7407407407407408
real_f1 = 0.7586206896551724
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 192.52it/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: 0.6160876750946045
step: 10, loss: 0.6018023490905762
step: 20, loss: 0.2681180238723755
step: 30, loss: 0.18086719512939453
step: 40, loss: 0.20037473738193512
step: 50, loss: 0.022347619757056236
step: 60, loss: 0.1854766309261322
step: 70, loss: 0.03433620184659958
step: 80, loss: 0.259749174118042
step: 90, loss: 0.11945196241140366
step: 100, loss: 0.0055139861069619656
step: 110, loss: 0.158479243516922
step: 120, loss: 0.007996385917067528
step: 130, loss: 0.010712778195738792
step: 140, loss: 0.00883543211966753
step: 150, loss: 0.03889644891023636
step: 160, loss: 0.006078546866774559
step: 170, loss: 0.1026257649064064
step: 180, loss: 0.006650817114859819
step: 190, loss: 0.012892743572592735
step: 200, loss: 0.024237196892499924
step: 210, loss: 0.008738347329199314
step: 220, loss: 0.008464175276458263
step: 230, loss: 0.03514161705970764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775784753363228, f1=0.9728506787330317,
best_f1=0.9728506787330317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013350460678339005
step: 10, loss: 0.01256899256259203
step: 20, loss: 0.14440888166427612
step: 30, loss: 0.192470520734787
step: 40, loss: 0.04191848263144493
step: 50, loss: 0.0142704788595438
step: 60, loss: 0.003815451869741082
step: 70, loss: 0.14706972241401672
step: 80, loss: 0.0030917120166122913
step: 90, loss: 0.029606731608510017
step: 100, loss: 0.003855952061712742
step: 110, loss: 0.03426450490951538
step: 120, loss: 0.017274700105190277
step: 130, loss: 0.003525871317833662
step: 140, loss: 0.0011079009855166078
step: 150, loss: 0.01482408121228218
step: 160, loss: 0.08440714329481125
step: 170, loss: 0.0029603014700114727
step: 180, loss: 0.0030776916537433863
step: 190, loss: 0.013415050692856312
step: 200, loss: 0.03348895534873009
step: 210, loss: 0.0012233792804181576
step: 220, loss: 0.12650133669376373
step: 230, loss: 0.11773496866226196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984090909090909, f1=0.9784824462061155,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034546179231256247
step: 10, loss: 0.0019857995212078094

```

```

step: 20, loss: 0.002006630878895521
step: 30, loss: 0.009160848334431648
step: 40, loss: 0.1411891132593155
step: 50, loss: 0.030618462711572647
step: 60, loss: 0.0033418352250009775
step: 70, loss: 0.002989760134369135
step: 80, loss: 0.0012401058338582516
step: 90, loss: 0.21989597380161285
step: 100, loss: 0.00045267361565493047
step: 110, loss: 0.008927907794713974
step: 120, loss: 0.015044140629470348
step: 130, loss: 0.002370523987337947
step: 140, loss: 0.0015473297098651528
step: 150, loss: 0.0011234706034883857
step: 160, loss: 0.01801779679954052
step: 170, loss: 0.003389124060049653
step: 180, loss: 0.003247613087296486
step: 190, loss: 0.001137851388193667
step: 200, loss: 0.04911498725414276
step: 210, loss: 0.0031491145491600037
step: 220, loss: 0.0011752460850402713
step: 230, loss: 0.002412821166217327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9852774631936579, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006021800800226629
step: 10, loss: 0.00032842805376276374
step: 20, loss: 0.0006343288696371019
step: 30, loss: 0.0006561180925928056
step: 40, loss: 0.0010310182115063071
step: 50, loss: 0.005246225744485855
step: 60, loss: 0.0020305230282247066
step: 70, loss: 0.0026736147701740265
step: 80, loss: 0.002126127015799284
step: 90, loss: 0.03878271207213402
step: 100, loss: 0.002600582782179117
step: 110, loss: 0.0015472408849745989
step: 120, loss: 0.024392282590270042
step: 130, loss: 0.013361495919525623

```

```

step: 140, loss: 0.002332124160602689
step: 150, loss: 0.11734918504953384
step: 160, loss: 0.0012736502103507519
step: 170, loss: 0.0066804904490709305
step: 180, loss: 0.00036191430990584195
step: 190, loss: 0.004251465201377869
step: 200, loss: 0.0045812674798071384
step: 210, loss: 0.01821957528591156
step: 220, loss: 0.00025382378953509033
step: 230, loss: 0.002504659118130803
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9775280898876404,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031720749102532864
step: 10, loss: 0.059211622923612595
step: 20, loss: 0.00046752826892770827
step: 30, loss: 0.00026547384914010763
step: 40, loss: 0.00023455399787053466
step: 50, loss: 0.00023505845456384122
step: 60, loss: 0.00017752127314452082
step: 70, loss: 0.0002850932942237705
step: 80, loss: 0.0003926422505173832
step: 90, loss: 0.00018988968804478645
step: 100, loss: 0.0020740048494189978
step: 110, loss: 0.0017462230753153563
step: 120, loss: 0.00045416128705255687
step: 130, loss: 0.0014475167263299227
step: 140, loss: 0.0010945504764094949
step: 150, loss: 0.0009551157709211111
step: 160, loss: 0.001842604367993772
step: 170, loss: 0.011562490835785866
step: 180, loss: 0.0039943503215909
step: 190, loss: 0.0009702492388896644
step: 200, loss: 0.0002972417860291898
step: 210, loss: 0.00022842800535727292
step: 220, loss: 0.00110739772208035
step: 230, loss: 0.0003450707299634814
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9863945578231292, f1=0.979591836734694,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027431712951511145
step: 10, loss: 0.0018098083091899753
step: 20, loss: 0.00181412766687572
step: 30, loss: 0.00035263318568468094
step: 40, loss: 0.00021490214567165822
step: 50, loss: 0.00024577201111242175
step: 60, loss: 0.0012251903535798192
step: 70, loss: 0.0021738784853368998
step: 80, loss: 0.002662754850462079
step: 90, loss: 0.014161123894155025
step: 100, loss: 0.026176149025559425
step: 110, loss: 0.00024485355243086815
step: 120, loss: 0.0005083243595436215
step: 130, loss: 0.004236895591020584
step: 140, loss: 0.0002534798695705831
step: 150, loss: 0.009849369525909424
step: 160, loss: 0.004498937167227268
step: 170, loss: 0.006775137037038803
step: 180, loss: 0.024405047297477722
step: 190, loss: 0.020158017054200172
step: 200, loss: 9.569044777890667e-05
step: 210, loss: 0.00016190327005460858
step: 220, loss: 0.07677457481622696
step: 230, loss: 0.008544377982616425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9808342728297633, f1=0.9841628959276018,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002067157533019781
step: 10, loss: 0.00022737204562872648
step: 20, loss: 0.0012219792697578669
step: 30, loss: 0.00042536744149401784
step: 40, loss: 0.0003855264512822032
step: 50, loss: 0.0003955896245315671
step: 60, loss: 0.00020934153872076422
step: 70, loss: 0.0005717657040804625
step: 80, loss: 0.00010340262087993324
step: 90, loss: 9.397843678016216e-05

```



```

step: 100, loss: 0.00010760100849438459
step: 110, loss: 0.0001267932530026883
step: 120, loss: 0.0003733652411028743
step: 130, loss: 0.0002303321671206504
step: 140, loss: 0.0012814165093004704
step: 150, loss: 0.008032613433897495
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.15215006470680237
step: 170, loss: 0.008186562918126583
step: 180, loss: 0.010567222721874714
step: 190, loss: 0.0022281284909695387
step: 200, loss: 0.07886621356010437
step: 210, loss: 7.202782580861822e-05
step: 220, loss: 0.0003333166823722422
step: 230, loss: 0.002163076773285866
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9785310734463276,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014635889965575188
step: 10, loss: 0.00022088034893386066
step: 20, loss: 0.00010885130177484825
step: 30, loss: 0.0001436168677173555
step: 40, loss: 0.00013158736692275852
step: 50, loss: 0.0016504063969478011
step: 60, loss: 6.149974069558084e-05
step: 70, loss: 0.00013633383787237108
step: 80, loss: 0.004310193005949259
step: 90, loss: 6.583526555914432e-05
step: 100, loss: 9.246812987839803e-05
step: 110, loss: 0.0002981809084303677
step: 120, loss: 0.0002575028338469565
step: 130, loss: 0.0001780326128937304
step: 140, loss: 9.68596141319722e-05
step: 150, loss: 0.00024810974719002843
step: 160, loss: 0.0001407427917001769
step: 170, loss: 0.00014777047908864915
step: 180, loss: 0.008218144066631794
step: 190, loss: 0.0011599333956837654
step: 200, loss: 0.0005101955612190068
step: 210, loss: 0.00018974574049934745
step: 220, loss: 0.00035344663774594665
step: 230, loss: 0.00011413659376557916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9875141884222476, f1=0.9773242630385486,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.317881681956351e-05
step: 10, loss: 0.001210426795296371
step: 20, loss: 0.0002301610802533105
step: 30, loss: 0.00025261801783926785
step: 40, loss: 0.04126520827412605
step: 50, loss: 7.118452049326152e-05
step: 60, loss: 0.00013722272706218064
step: 70, loss: 0.0004085826512891799
step: 80, loss: 0.000113921363663394
step: 90, loss: 0.00013491125719156116
step: 100, loss: 0.0002860722888726741
step: 110, loss: 0.00014891741739120334
step: 120, loss: 0.0003106636868324131
step: 130, loss: 7.169425953179598e-05
step: 140, loss: 7.975712651386857e-05
step: 150, loss: 0.0009823720902204514
step: 160, loss: 0.00011240984167670831
step: 170, loss: 0.00018419271509628743
step: 180, loss: 0.003946591634303331
step: 190, loss: 0.0031248906161636114
step: 200, loss: 0.00016765334294177592
step: 210, loss: 0.0010742044541984797
step: 220, loss: 0.0005555255338549614
step: 230, loss: 0.00017454901535529643
/usr/local/lib/python3.7/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=0.9796380090497738, f1=0.9774266365688488,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003403273585718125
step: 10, loss: 0.0001025576566462405
step: 20, loss: 5.247506487648934e-05
step: 30, loss: 0.0001529433357063681
step: 40, loss: 7.457844185410067e-05

```

```

step: 50, loss: 0.000928123074118048
step: 60, loss: 0.0007978428620845079
step: 70, loss: 0.0006757571827620268
step: 80, loss: 0.00016458617756143212
step: 90, loss: 0.0005120630958117545
step: 100, loss: 0.0007627706509083509
step: 110, loss: 0.00031156474142335355
step: 120, loss: 0.0012099641608074307
step: 130, loss: 0.00014727719826623797
step: 140, loss: 0.0003940555907320231
step: 150, loss: 0.08449140191078186
step: 160, loss: 7.379780436167493e-05
step: 170, loss: 0.0002692803682293743
step: 180, loss: 0.0005908907041884959
step: 190, loss: 0.01817174442112446
step: 200, loss: 8.263201743829995e-05
step: 210, loss: 7.676012319279835e-05
step: 220, loss: 0.00015503317990805954
step: 230, loss: 0.00013328998466022313
/usr/local/lib/python3.7/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=0.9875141884222476, f1=0.9774266365688488,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.858639269717969e-05
step: 10, loss: 0.001845458522439003
step: 20, loss: 0.0025268886238336563
step: 30, loss: 0.004521454684436321
step: 40, loss: 0.0010890841949731112
step: 50, loss: 0.0006638956838287413
step: 60, loss: 9.751596371643245e-05
step: 70, loss: 0.00046564877266064286
step: 80, loss: 4.356162025942467e-05
step: 90, loss: 0.00015100683958735317
step: 100, loss: 6.712022877763957e-05
step: 110, loss: 0.0001410567492712289
step: 120, loss: 7.046532118692994e-05
step: 130, loss: 0.0006261744420044124
step: 140, loss: 6.798693357268348e-05
step: 150, loss: 0.035053592175245285
step: 160, loss: 6.114948337199166e-05
step: 170, loss: 0.015075819566845894
step: 180, loss: 4.2794425098691136e-05
step: 190, loss: 4.172561602899805e-05
step: 200, loss: 6.41128426650539e-05

```

```

step: 210, loss: 9.949040395440534e-05
step: 220, loss: 4.4519292714539915e-05
step: 230, loss: 0.0001706875627860427
/usr/local/lib/python3.7/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=0.9853107344632768, f1=0.9841986455981941,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.081039489596151e-05
step: 10, loss: 0.0002925916633103043
step: 20, loss: 4.89172525703907e-05
step: 30, loss: 0.00012835251982323825
step: 40, loss: 0.0001252399815712124
step: 50, loss: 0.00018556472787167877
step: 60, loss: 0.00034460038295947015
step: 70, loss: 9.98566611087881e-05
step: 80, loss: 0.005743478424847126
step: 90, loss: 2.6404373784316704e-05
step: 100, loss: 0.00041355082066729665
step: 110, loss: 5.85090383538045e-05
step: 120, loss: 0.0003947024524677545
step: 130, loss: 3.478218422969803e-05
step: 140, loss: 3.733365883817896e-05
step: 150, loss: 4.730048021883704e-05
step: 160, loss: 3.7009172956459224e-05
step: 170, loss: 9.293061884818599e-05
step: 180, loss: 0.00019319847342558205
step: 190, loss: 4.5019878598395735e-05
step: 200, loss: 2.7134419724461623e-05
step: 210, loss: 4.037998951389454e-05
step: 220, loss: 3.424156966502778e-05
step: 230, loss: 0.052252013236284256
/usr/local/lib/python3.7/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=0.9852104664391355, f1=0.9807037457434733,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016174526535905898
step: 10, loss: 0.00019348332716617733
step: 20, loss: 0.00035463995300233364
step: 30, loss: 5.231301474850625e-05
step: 40, loss: 0.0003660500515252352

```

```

step: 50, loss: 4.4865642848890275e-05
step: 60, loss: 3.404817107366398e-05
step: 70, loss: 2.5089359041885473e-05
step: 80, loss: 2.7741694793803617e-05
step: 90, loss: 6.884805770823732e-05
step: 100, loss: 2.8098986149416305e-05
step: 110, loss: 0.024962354451417923
step: 120, loss: 0.03659442812204361
step: 130, loss: 5.7144261518260464e-05
step: 140, loss: 2.7737927666748874e-05
step: 150, loss: 2.9663857276318595e-05
step: 160, loss: 5.152618541615084e-05
step: 170, loss: 8.078755490714684e-05
step: 180, loss: 7.237375393742695e-05
step: 190, loss: 4.8345260438509285e-05
step: 200, loss: 0.0010405739303678274
step: 210, loss: 2.771107028820552e-05
step: 220, loss: 2.657563891261816e-05
step: 230, loss: 3.53317336703185e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9852440408626559, f1=0.9807909604519773,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.580085492809303e-05
step: 10, loss: 0.0005257012671791017
step: 20, loss: 3.893667962984182e-05
step: 30, loss: 3.542244667187333e-05
step: 40, loss: 4.293794700060971e-05
step: 50, loss: 4.139192242291756e-05
step: 60, loss: 3.40472579409834e-05
step: 70, loss: 4.2949912312906235e-05
step: 80, loss: 2.6903468096861616e-05
step: 90, loss: 3.409264172660187e-05
step: 100, loss: 3.6356665077619255e-05
step: 110, loss: 2.9517388611566275e-05
step: 120, loss: 1.8123322661267594e-05
step: 130, loss: 2.9134589567547664e-05
step: 140, loss: 3.6152279790258035e-05
step: 150, loss: 5.972018334432505e-05
step: 160, loss: 0.0071009136736392975
step: 170, loss: 2.383026367169805e-05
step: 180, loss: 4.121972597204149e-05
step: 190, loss: 4.500124850892462e-05
step: 200, loss: 3.674238905659877e-05

```

```

step: 210, loss: 3.9631922845728695e-05
step: 220, loss: 4.2728577682282776e-05
step: 230, loss: 8.827705460134894e-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=0.9841269841269841, f1=0.9796839729119639,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.48508840741124e-05
step: 10, loss: 2.264534487039782e-05
step: 20, loss: 7.261910650413483e-05
step: 30, loss: 3.9247544918907806e-05
step: 40, loss: 5.6631128245498985e-05
step: 50, loss: 0.0034759780392050743
step: 60, loss: 5.1198814617237076e-05
step: 70, loss: 3.550750625436194e-05
step: 80, loss: 5.660008901031688e-05
step: 90, loss: 8.094956137938425e-05
step: 100, loss: 3.5995402868138626e-05
step: 110, loss: 3.7084850191604346e-05
step: 120, loss: 3.775065124500543e-05
step: 130, loss: 4.912374424748123e-05
step: 140, loss: 2.2474321667687036e-05
step: 150, loss: 9.206754475599155e-05
step: 160, loss: 3.2962983823381364e-05
step: 170, loss: 2.614327604533173e-05
step: 180, loss: 0.0001458519254811108
step: 190, loss: 3.3101707231253386e-05
step: 200, loss: 8.670039096614346e-05
step: 210, loss: 2.844554910552688e-05
step: 220, loss: 6.036358172423206e-05
step: 230, loss: 2.5927518436219543e-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=0.9841269841269841, f1=0.9785310734463276,
best_f1=0.9773242630385486
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled               : True
opt_level             : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:15, 160.90it/s]
load_f1 = 0.9829738933030647
real_f1 = 0.9807474518686297
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.83it/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, 289kB/s]
 Downloading: 100% 232k/232k [00:00<00:00, 256kB/s]
 Downloading: 100% 440M/440M [00:08<00:00, 54.5MB/s]
 Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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


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

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

```
step: 0, loss: 0.6131656765937805
step: 10, loss: 0.5458610653877258
step: 20, loss: 0.5058370232582092
step: 30, loss: 0.15447235107421875
step: 40, loss: 0.14685778319835663
step: 50, loss: 0.1476902961730957
step: 60, loss: 0.04248557612299919
step: 70, loss: 0.11163368821144104
step: 80, loss: 0.0776122435927391
step: 90, loss: 0.17099924385547638
step: 100, loss: 0.0525943785905838
step: 110, loss: 0.09751500934362411
step: 120, loss: 0.0867849811911583
step: 130, loss: 0.04742724448442459
step: 140, loss: 0.25287553668022156
step: 150, loss: 0.04952855780720711
step: 160, loss: 0.061722543090581894
step: 170, loss: 0.22188138961791992
step: 180, loss: 0.18582426011562347
step: 190, loss: 0.02190866693854332
step: 200, loss: 0.1433989703655243
step: 210, loss: 0.16660095751285553
step: 220, loss: 0.24091005325317383
step: 230, loss: 0.13381385803222656
step: 240, loss: 0.21433721482753754
step: 250, loss: 0.028790626674890518
step: 260, loss: 0.05066675692796707
step: 270, loss: 0.02707209624350071
step: 280, loss: 0.111310675740242
step: 290, loss: 0.056364450603723526
step: 300, loss: 0.042654287070035934
step: 310, loss: 0.2655971646308899
step: 320, loss: 0.10262089222669601
step: 330, loss: 0.0629330649971962
step: 340, loss: 0.020097561180591583
step: 350, loss: 0.15459631383419037
step: 360, loss: 0.05065270885825157
step: 370, loss: 0.06070220097899437
step: 380, loss: 0.05613747611641884
```

```

step: 390, loss: 0.11426419764757156
step: 400, loss: 0.21832701563835144
step: 410, loss: 0.0588616207242012
step: 420, loss: 0.02027837559580803
step: 430, loss: 0.17252127826213837
step: 440, loss: 0.018073372542858124
step: 450, loss: 0.01458664145320654
step: 460, loss: 0.008270681835711002
step: 470, loss: 0.19439199566841125
step: 480, loss: 0.08142128586769104
step: 490, loss: 0.1538284420967102
step: 500, loss: 0.03482329845428467
step: 510, loss: 0.1016848236322403
step: 520, loss: 0.06757648289203644
step: 530, loss: 0.005553641356527805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9234972677595629, f1=0.913457181694608,
best_f1=0.913457181694608
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1305985152721405
step: 10, loss: 0.030911823734641075
step: 20, loss: 0.04126792773604393
step: 30, loss: 0.10978595167398453
step: 40, loss: 0.12300001084804535
step: 50, loss: 0.21756966412067413
step: 60, loss: 0.024722937494516373
step: 70, loss: 0.01998637057840824
step: 80, loss: 0.07076039910316467
step: 90, loss: 0.015235917642712593
step: 100, loss: 0.0680018737912178
step: 110, loss: 0.043319132179021835
step: 120, loss: 0.10190130770206451
step: 130, loss: 0.18661358952522278
step: 140, loss: 0.027722137048840523
step: 150, loss: 0.05659234896302223
step: 160, loss: 0.07032795250415802
step: 170, loss: 0.025896186009049416
step: 180, loss: 0.037320010364055634
step: 190, loss: 0.09087035804986954
step: 200, loss: 0.01340766903012991

```

```

step: 210, loss: 0.02226976305246353
step: 220, loss: 0.08735481649637222
step: 230, loss: 0.006780026480555534
step: 240, loss: 0.06627029925584793
step: 250, loss: 0.002735364018008113
step: 260, loss: 0.0056840842589735985
step: 270, loss: 0.3281061351299286
step: 280, loss: 0.026774635538458824
step: 290, loss: 0.039085179567337036
step: 300, loss: 0.18705812096595764
step: 310, loss: 0.020771857351064682
step: 320, loss: 0.04258488118648529
step: 330, loss: 0.03876735642552376
step: 340, loss: 0.01081305742263794
step: 350, loss: 0.172308549284935
step: 360, loss: 0.19582411646842957
step: 370, loss: 0.12950780987739563
step: 380, loss: 0.1112857311964035
step: 390, loss: 0.07132012397050858
step: 400, loss: 0.054627515375614166
step: 410, loss: 0.022618787363171577
step: 420, loss: 0.08627623319625854
step: 430, loss: 0.003754950128495693
step: 440, loss: 0.1264556348323822
step: 450, loss: 0.007657757960259914
step: 460, loss: 0.028602948412299156
step: 470, loss: 0.14125174283981323
step: 480, loss: 0.22733300924301147
step: 490, loss: 0.013397961854934692
step: 500, loss: 0.25262150168418884
step: 510, loss: 0.017906788736581802
step: 520, loss: 0.021194148808717728
step: 530, loss: 0.031111840158700943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9266266728195662, f1=0.926605504587156,
best_f1=0.926605504587156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02678108774125576
step: 10, loss: 0.046760592609643936
step: 20, loss: 0.07594453543424606

```

step: 30, loss: 0.06872236728668213
step: 40, loss: 0.0038001397624611855
step: 50, loss: 0.1758647859096527
step: 60, loss: 0.014343290589749813
step: 70, loss: 0.004514650907367468
step: 80, loss: 0.0014125462621450424
step: 90, loss: 0.009315931238234043
step: 100, loss: 0.03360210731625557
step: 110, loss: 0.00443840678781271
step: 120, loss: 0.033248938620090485
step: 130, loss: 0.09761621057987213
step: 140, loss: 0.0729832723736763
step: 150, loss: 0.0314851738512516
step: 160, loss: 0.03203263506293297
step: 170, loss: 0.03640840947628021
step: 180, loss: 0.036208897829055786
step: 190, loss: 0.004420209676027298
step: 200, loss: 0.013377116061747074
step: 210, loss: 0.04865632951259613
step: 220, loss: 0.11405551433563232
step: 230, loss: 0.18915890157222748
step: 240, loss: 0.003535299561917782
step: 250, loss: 0.026771824806928635
step: 260, loss: 0.032778050750494
step: 270, loss: 0.007490411400794983
step: 280, loss: 0.18551599979400635
step: 290, loss: 0.006641969084739685
step: 300, loss: 0.1315172165632248
step: 310, loss: 0.01085018552839756
step: 320, loss: 0.009458638727664948
step: 330, loss: 0.004372093360871077
step: 340, loss: 0.017262034118175507
step: 350, loss: 0.008621478453278542
step: 360, loss: 0.03596022352576256
step: 370, loss: 0.008347981609404087
step: 380, loss: 0.003852132009342313
step: 390, loss: 0.035826053470373154
step: 400, loss: 0.055689338594675064
step: 410, loss: 0.005416004452854395
step: 420, loss: 0.1468142718076706
step: 430, loss: 0.06069837510585785
step: 440, loss: 0.0029892357997596264
step: 450, loss: 0.08638142049312592
step: 460, loss: 0.018461931496858597
step: 470, loss: 0.04480127617716789
step: 480, loss: 0.006850267760455608
step: 490, loss: 0.005469798110425472
step: 500, loss: 0.0024734449107199907

```

step: 510, loss: 0.006273058243095875
step: 520, loss: 0.030196253210306168
step: 530, loss: 0.027587831020355225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.927710843373494, f1=0.9285384970032272,
best_f1=0.9285384970032272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014804691076278687
step: 10, loss: 0.012086433358490467
step: 20, loss: 0.020868618041276932
step: 30, loss: 0.005080789793282747
step: 40, loss: 0.02273429185152054
step: 50, loss: 0.012821373529732227
step: 60, loss: 0.020252780988812447
step: 70, loss: 0.0015752790495753288
step: 80, loss: 0.010092348791658878
step: 90, loss: 0.010008088313043118
step: 100, loss: 0.0009223191882483661
step: 110, loss: 0.013730574399232864
step: 120, loss: 0.001716467086225748
step: 130, loss: 0.00043538110912777483
step: 140, loss: 0.0006742006517015398
step: 150, loss: 0.009345073252916336
step: 160, loss: 0.04001487046480179
step: 170, loss: 0.006178424693644047
step: 180, loss: 0.004946308210492134
step: 190, loss: 0.00573036540299654
step: 200, loss: 0.0030443419236689806
step: 210, loss: 0.020636631175875664
step: 220, loss: 0.002836846746504307
step: 230, loss: 0.06946935504674911
step: 240, loss: 0.005375944077968597
step: 250, loss: 0.019950436428189278
step: 260, loss: 0.005844524595886469
step: 270, loss: 0.005779662169516087
step: 280, loss: 0.017766520380973816
step: 290, loss: 0.06899330019950867
step: 300, loss: 0.0047972677275538445
step: 310, loss: 0.024105791002511978
step: 320, loss: 0.0941750556230545

```

```

step: 330, loss: 0.15647543966770172
step: 340, loss: 0.007590575609356165
step: 350, loss: 0.006063947454094887
step: 360, loss: 0.00699726864695549
step: 370, loss: 0.004108448047190905
step: 380, loss: 0.0012377570383250713
step: 390, loss: 0.00087932770838961
step: 400, loss: 0.20908403396606445
step: 410, loss: 0.0013914033770561218
step: 420, loss: 0.06749977916479111
step: 430, loss: 0.08145041763782501
step: 440, loss: 0.0013273631921038032
step: 450, loss: 0.001638802932575345
step: 460, loss: 0.03902290388941765
step: 470, loss: 0.0019619956146925688
step: 480, loss: 0.0287595484405756
step: 490, loss: 0.0032333971466869116
step: 500, loss: 0.010242629796266556
step: 510, loss: 0.020034845918416977
step: 520, loss: 0.11103448271751404
step: 530, loss: 0.014600371941924095
/usr/local/lib/python3.7/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=0.9255966307908282, f1=0.925497454881999,
best_f1=0.9285384970032272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043291281908750534
step: 10, loss: 0.12854821979999542
step: 20, loss: 0.0019938405603170395
step: 30, loss: 0.0057868813164532185
step: 40, loss: 0.03703327476978302
step: 50, loss: 0.006276112515479326
step: 60, loss: 0.15512603521347046
step: 70, loss: 0.019740235060453415
step: 80, loss: 0.0008248545927926898
step: 90, loss: 0.020496107637882233
step: 100, loss: 0.03170965611934662
step: 110, loss: 0.0004636347875930369
step: 120, loss: 0.034992434084415436
step: 130, loss: 0.0007003091741353273
step: 140, loss: 0.0007035773014649749
step: 150, loss: 0.02567366510629654
step: 160, loss: 0.00047185615403577685
step: 170, loss: 0.014311667531728745
step: 180, loss: 0.0010864053620025516

```

```

step: 190, loss: 0.00121020816732198
step: 200, loss: 0.0013233856298029423
step: 210, loss: 0.001415578997693956
step: 220, loss: 0.009321454912424088
step: 230, loss: 0.0013668567407876253
step: 240, loss: 0.003309893887490034
step: 250, loss: 0.0005285277147777379
step: 260, loss: 0.002256567357107997
step: 270, loss: 0.009686905890703201
step: 280, loss: 0.008804845623672009
step: 290, loss: 0.12971149384975433
step: 300, loss: 0.0015979389427229762
step: 310, loss: 0.003377100918442011
step: 320, loss: 0.1381145864725113
step: 330, loss: 0.011835281737148762
step: 340, loss: 0.05409636348485947
step: 350, loss: 0.0023051791358739138
step: 360, loss: 0.002527270233258605
step: 370, loss: 0.0032353391870856285
step: 380, loss: 0.0014200415462255478
step: 390, loss: 0.00026720380992628634
step: 400, loss: 0.00825646985322237
step: 410, loss: 0.0012494797119870782
step: 420, loss: 0.012193293310701847
step: 430, loss: 0.0008605736657045782
step: 440, loss: 0.10051179677248001
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 450, loss: 0.1461179107427597
step: 460, loss: 0.005174926016479731
step: 470, loss: 0.0035996234510093927
step: 480, loss: 0.016104472801089287
step: 490, loss: 0.0006035795086063445
step: 500, loss: 0.007561319041997194
step: 510, loss: 0.022437019273638725
step: 520, loss: 0.006287725176662207
step: 530, loss: 0.0028350530192255974
/usr/local/lib/python3.7/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=0.9186875891583451, f1=0.9164700330344502,
best_f1=0.9285384970032272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004058316349983215
step: 10, loss: 0.0001838245225371793
step: 20, loss: 0.040564149618148804
step: 30, loss: 0.0011987807229161263

```

step: 40, loss: 0.05114335939288139
step: 50, loss: 0.07196895033121109
step: 60, loss: 0.00016412034165114164
step: 70, loss: 0.00013274408411234617
step: 80, loss: 0.0007327237399294972
step: 90, loss: 0.010106394998729229
step: 100, loss: 0.0016318743582814932
step: 110, loss: 0.0010711518116295338
step: 120, loss: 0.0010127925779670477
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.0013157054781913757
step: 140, loss: 0.021876949816942215
step: 150, loss: 0.0016078135231509805
step: 160, loss: 0.0013090481515973806
step: 170, loss: 0.0006438368000090122
step: 180, loss: 0.0003600372583605349
step: 190, loss: 0.00040750522748567164
step: 200, loss: 0.0005405033007264137
step: 210, loss: 0.01024746522307396
step: 220, loss: 0.007990465499460697
step: 230, loss: 0.002026102039963007
step: 240, loss: 0.039435695856809616
step: 250, loss: 0.00041031098226085305
step: 260, loss: 0.00019969682034570724
step: 270, loss: 0.049291349947452545
step: 280, loss: 0.0019438094459474087
step: 290, loss: 0.004202405922114849
step: 300, loss: 0.0016126229893416166
step: 310, loss: 0.0010413707932457328
step: 320, loss: 0.0005364437238313258
step: 330, loss: 0.002188070910051465
step: 340, loss: 0.02774086222052574
step: 350, loss: 0.0023191431537270546
step: 360, loss: 0.003170965239405632
step: 370, loss: 0.003434722777456045
step: 380, loss: 0.00022205521236173809
step: 390, loss: 0.01229170709848404
step: 400, loss: 0.00012634861923288554
step: 410, loss: 0.00032810488482937217
step: 420, loss: 0.0011282290797680616
step: 430, loss: 0.00041956492350436747
step: 440, loss: 0.00023225814220495522
step: 450, loss: 0.001395867089740932
step: 460, loss: 9.816299279918894e-05
step: 470, loss: 0.007431095466017723
step: 480, loss: 0.003999523818492889
step: 490, loss: 0.018637122586369514
step: 500, loss: 0.0012754888739436865


```

step: 510, loss: 0.003633993212133646
step: 520, loss: 0.008581340312957764
step: 530, loss: 0.010082930326461792
/usr/local/lib/python3.7/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=0.9163498098859316, f1=0.9202626641651033,
best_f1=0.9285384970032272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021738703071605414
step: 10, loss: 0.0066016665659844875
step: 20, loss: 0.004434582777321339
step: 30, loss: 0.0007183939451351762
step: 40, loss: 0.0037869391962885857
step: 50, loss: 0.015510812401771545
step: 60, loss: 0.0012470022775232792
step: 70, loss: 0.007227855734527111
step: 80, loss: 0.0031273963395506144
step: 90, loss: 0.00048213504487648606
step: 100, loss: 0.00029935178463347256
step: 110, loss: 0.011442124843597412
step: 120, loss: 0.00025160194491036236
step: 130, loss: 0.0005743289948441088
step: 140, loss: 0.0004558709042612463
step: 150, loss: 0.003920306917279959
step: 160, loss: 0.0007167247822508216
step: 170, loss: 0.0010188949527218938
step: 180, loss: 0.0018948856741189957
step: 190, loss: 0.021148094907402992
step: 200, loss: 0.00026757948216982186
step: 210, loss: 0.02699396014213562
step: 220, loss: 0.0009421081631444395
step: 230, loss: 0.0068640802055597305
step: 240, loss: 0.0011676347348839045
step: 250, loss: 0.0011712070554494858
step: 260, loss: 0.00027767583378590643
step: 270, loss: 0.0004405493091326207
step: 280, loss: 0.016141362488269806
step: 290, loss: 0.0017366843530908227
step: 300, loss: 0.004615856800228357
step: 310, loss: 0.0003020838485099375
step: 320, loss: 9.190929995384067e-05
step: 330, loss: 0.0001463072549086064
step: 340, loss: 0.004520641174167395
step: 350, loss: 0.00040708479355089366
step: 360, loss: 0.0006666618282906711

```

```

step: 370, loss: 0.0005783864762634039
step: 380, loss: 0.00375442230142653
step: 390, loss: 0.00017146665777545422
step: 400, loss: 0.00423762621358037
step: 410, loss: 0.0007924732053652406
step: 420, loss: 0.007538542151451111
step: 430, loss: 0.006341320462524891
step: 440, loss: 0.0007803954067640007
step: 450, loss: 0.026004116982221603
step: 460, loss: 0.0002460702962707728
step: 470, loss: 0.09539392590522766
step: 480, loss: 0.16107897460460663
step: 490, loss: 0.0028791569638997316
step: 500, loss: 0.0009505567140877247
step: 510, loss: 0.0027590489480644464
step: 520, loss: 0.006974717602133751
step: 530, loss: 0.003945984877645969
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9331489165514063, f1=0.9222988505747127,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010921044740825891
step: 10, loss: 0.020361388102173805
step: 20, loss: 0.08378364145755768
step: 30, loss: 0.008279534988105297
step: 40, loss: 0.00034030849928967655
step: 50, loss: 0.005392452701926231
step: 60, loss: 0.0024112381506711245
step: 70, loss: 0.0002028268645517528
step: 80, loss: 0.011014261282980442
step: 90, loss: 0.0009648966370150447
step: 100, loss: 0.00010070623102365062
step: 110, loss: 0.0002112688816851005
step: 120, loss: 0.0013201559195294976
step: 130, loss: 0.000734530040062964
step: 140, loss: 0.00029166246531531215
step: 150, loss: 0.00012841868738178164
step: 160, loss: 0.024322422221302986
step: 170, loss: 0.00032959316740743816
step: 180, loss: 0.0003277207724750042

```

```

step: 190, loss: 0.003552323207259178
step: 200, loss: 0.00249241734854877
step: 210, loss: 0.00034394761314615607
step: 220, loss: 0.0009527344373054802
step: 230, loss: 0.0052878460846841335
step: 240, loss: 0.00012086041533621028
step: 250, loss: 3.9583377656526864e-05
step: 260, loss: 0.0007245094748213887
step: 270, loss: 0.006818788591772318
step: 280, loss: 0.008596500381827354
step: 290, loss: 6.44441315671429e-05
step: 300, loss: 0.0010301503352820873
step: 310, loss: 0.009022170677781105
step: 320, loss: 0.0010068707633763552
step: 330, loss: 0.0003560494224075228
step: 340, loss: 0.0006923438631929457
step: 350, loss: 0.0003758264647331089
step: 360, loss: 0.010658439248800278
step: 370, loss: 0.0002450503525324166
step: 380, loss: 0.0011224672198295593
step: 390, loss: 0.13025808334350586
step: 400, loss: 0.000171125924680382
step: 410, loss: 0.00013490533456206322
step: 420, loss: 0.010449953377246857
step: 430, loss: 0.011853385716676712
step: 440, loss: 0.07594014704227448
step: 450, loss: 0.0006051716045476496
step: 460, loss: 0.00031532542197965086
step: 470, loss: 7.018230098765343e-05
step: 480, loss: 5.0195918447570875e-05
step: 490, loss: 0.0005498963291756809
step: 500, loss: 7.722756708972156e-05
step: 510, loss: 6.876578117953613e-05
step: 520, loss: 0.0035827327519655228
step: 530, loss: 9.892028901958838e-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=0.928772258669166, f1=0.9211502782931354,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010357258142903447
step: 10, loss: 0.0009734708000905812
step: 20, loss: 0.01923714205622673
step: 30, loss: 0.00017981410201173276
step: 40, loss: 0.0017274554120376706

```

step: 50, loss: 0.0001310365623794496
step: 60, loss: 0.0003583342186175287
step: 70, loss: 0.0008943747379817069
step: 80, loss: 0.0007519164937548339
step: 90, loss: 0.0005022430559620261
step: 100, loss: 0.00014042614202480763
step: 110, loss: 0.00015568580420222133
step: 120, loss: 0.00027589488308876753
step: 130, loss: 0.0016645489959046245
step: 140, loss: 0.0007608064915984869
step: 150, loss: 0.00015988202358130366
step: 160, loss: 0.0004945645923726261
step: 170, loss: 0.006144167389720678
step: 180, loss: 0.00040157182957045734
step: 190, loss: 0.00791709590703249
step: 200, loss: 8.192509994842112e-05
step: 210, loss: 0.1543179452419281
step: 220, loss: 0.0014282730408012867
step: 230, loss: 0.02345952019095421
step: 240, loss: 0.0019041035557165742
step: 250, loss: 0.00032100133830681443
step: 260, loss: 0.004358056001365185
step: 270, loss: 5.694273568224162e-05
step: 280, loss: 0.0001923874660860747
step: 290, loss: 0.04751084744930267
step: 300, loss: 8.211174281314015e-05
step: 310, loss: 0.015100072138011456
step: 320, loss: 0.00026588627952151
step: 330, loss: 0.00325647066347301
step: 340, loss: 0.0005455795908346772
step: 350, loss: 0.00010868873505387455
step: 360, loss: 3.3891825296450406e-05
step: 370, loss: 0.0004447197716217488
step: 380, loss: 0.00011239347804803401
step: 390, loss: 6.765041325706989e-05
step: 400, loss: 0.03620930761098862
step: 410, loss: 0.0011681760661303997
step: 420, loss: 4.565929339150898e-05
step: 430, loss: 0.00017882413521874696
step: 440, loss: 0.0011872949544340372
step: 450, loss: 0.00010721817670855671
step: 460, loss: 7.978601934155449e-05
step: 470, loss: 8.555312524549663e-05
step: 480, loss: 5.0915623432956636e-05
step: 490, loss: 0.00013587398279923946
step: 500, loss: 0.0011068281019106507
step: 510, loss: 0.0004409915709402412
step: 520, loss: 0.0008438042714260519

```

step: 530, loss: 0.0003293339104857296
/usr/local/lib/python3.7/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=0.9273323956868261, f1=0.9265799256505577,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015680679643992335
step: 10, loss: 4.828748933505267e-05
step: 20, loss: 0.00012043834431096911
step: 30, loss: 9.270678856410086e-05
step: 40, loss: 0.0007738824933767319
step: 50, loss: 0.00010209157335339114
step: 60, loss: 3.770979310502298e-05
step: 70, loss: 6.449321517720819e-05
step: 80, loss: 3.9322763768723235e-05
step: 90, loss: 4.470790736377239e-05
step: 100, loss: 3.734450729098171e-05
step: 110, loss: 3.833141818176955e-05
step: 120, loss: 0.00021239471971057355
step: 130, loss: 3.9337497582891956e-05
step: 140, loss: 0.007993093691766262
step: 150, loss: 4.580306267598644e-05
step: 160, loss: 0.0007461826899088919
step: 170, loss: 7.867729436839e-05
step: 180, loss: 6.659035716438666e-05
step: 190, loss: 0.00011180675210198388
step: 200, loss: 0.00034112768480554223
step: 210, loss: 4.933813397656195e-05
step: 220, loss: 0.0009684744873084128
step: 230, loss: 0.0001458650513086468
step: 240, loss: 4.7137804358499125e-05
step: 250, loss: 0.0018064940813928843
step: 260, loss: 0.00031190848676487803
step: 270, loss: 0.00022566519328393042
step: 280, loss: 0.0009958567097783089
step: 290, loss: 4.39781833847519e-05
step: 300, loss: 0.000292174139758572
step: 310, loss: 0.00019844478811137378
step: 320, loss: 0.006015370134264231
step: 330, loss: 0.07953262329101562
step: 340, loss: 0.11812973022460938
step: 350, loss: 0.23891304433345795
step: 360, loss: 0.0005034773494116962
step: 370, loss: 0.003968555014580488
step: 380, loss: 0.00373770366422832

```

```

step: 390, loss: 0.010802504606544971
step: 400, loss: 0.0008199251024052501
step: 410, loss: 0.003184306202456355
step: 420, loss: 0.00035524129634723067
step: 430, loss: 0.0004685016756411642
step: 440, loss: 0.0023057428188622
step: 450, loss: 8.31131765153259e-05
step: 460, loss: 6.49130015517585e-05
step: 470, loss: 0.00021413227659650147
step: 480, loss: 9.14420306799002e-05
step: 490, loss: 0.0009731760947033763
step: 500, loss: 0.004903409630060196
step: 510, loss: 5.386117845773697e-05
step: 520, loss: 8.749881817493588e-05
step: 530, loss: 0.0010404884815216064
/usr/local/lib/python3.7/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=0.9322191272051997, f1=0.9294605809128631,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009660134674049914
step: 10, loss: 7.21010728739202e-05
step: 20, loss: 0.00011268431262578815
step: 30, loss: 0.0019629704765975475
step: 40, loss: 4.33086788689252e-05
step: 50, loss: 0.000813081453088671
step: 60, loss: 0.0035864838864654303
step: 70, loss: 0.00016378966392949224
step: 80, loss: 0.0077219558879733086
step: 90, loss: 4.50500774604734e-05
step: 100, loss: 0.0038780171889811754
step: 110, loss: 0.00016249463078565896
step: 120, loss: 0.0007657759706489742
step: 130, loss: 0.00014982848369982094
step: 140, loss: 0.00013913247676100582
step: 150, loss: 2.6989153411705047e-05
step: 160, loss: 0.0004889483097940683
step: 170, loss: 4.342377360444516e-05
step: 180, loss: 6.835486419731751e-05
step: 190, loss: 0.00027528416831046343
step: 200, loss: 0.0021802678238600492
step: 210, loss: 0.00018617989553604275
step: 220, loss: 0.00046130173723213375
step: 230, loss: 2.9108703529345803e-05
step: 240, loss: 0.0001407690579071641

```

```

step: 250, loss: 0.0028639452066272497
step: 260, loss: 2.988718006236013e-05
step: 270, loss: 0.0003528591769281775
step: 280, loss: 0.0010520886862650514
step: 290, loss: 6.482327444246039e-05
step: 300, loss: 0.0002893354103434831
step: 310, loss: 0.00016966431576292962
step: 320, loss: 0.0001957167114596814
step: 330, loss: 6.306588329607621e-05
step: 340, loss: 6.378210673574358e-05
step: 350, loss: 8.44752139528282e-05
step: 360, loss: 0.00020311646221671253
step: 370, loss: 0.00026261014863848686
step: 380, loss: 4.328084469307214e-05
step: 390, loss: 3.9991569792618975e-05
step: 400, loss: 0.0026106329169124365
step: 410, loss: 3.5232405934948474e-05
step: 420, loss: 4.042144792038016e-05
step: 430, loss: 4.159224044997245e-05
step: 440, loss: 0.00016772831440903246
step: 450, loss: 0.00046256970381364226
step: 460, loss: 0.0006466695922426879
step: 470, loss: 9.457465057494119e-05
step: 480, loss: 0.00011145022290293127
step: 490, loss: 0.0004910605493932962
step: 500, loss: 0.0010309139033779502
step: 510, loss: 0.0002679661847651005
step: 520, loss: 2.8020938771078363e-05
step: 530, loss: 3.789913171203807e-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=0.9279999999999999, f1=0.9207089552238805,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.7651072489097714e-05
step: 10, loss: 2.9238393835839815e-05
step: 20, loss: 0.00010491560533409938
step: 30, loss: 0.0017742933705449104
step: 40, loss: 0.0002698609314393252
step: 50, loss: 0.03850200027227402
step: 60, loss: 0.0014946392038837075
step: 70, loss: 5.503812280949205e-05
step: 80, loss: 3.321683834656142e-05
step: 90, loss: 0.001357845263555646
step: 100, loss: 0.0005885578575544059

```

```
step: 110, loss: 3.770927287405357e-05
step: 120, loss: 5.4235290008364245e-05
step: 130, loss: 0.00032224980532191694
step: 140, loss: 7.719272252870724e-05
step: 150, loss: 3.876471964758821e-05
step: 160, loss: 2.099907760566566e-05
step: 170, loss: 4.816586442757398e-05
step: 180, loss: 6.767593731638044e-05
step: 190, loss: 3.242734965169802e-05
step: 200, loss: 0.0006487223436124623
step: 210, loss: 3.386849130038172e-05
step: 220, loss: 6.027620838722214e-05
step: 230, loss: 1.6938689441303723e-05
step: 240, loss: 5.0249233026988804e-05
step: 250, loss: 5.944000440649688e-05
step: 260, loss: 2.7089401555713266e-05
step: 270, loss: 3.651745282695629e-05
step: 280, loss: 0.0010484351078048348
step: 290, loss: 0.00027757376665249467
step: 300, loss: 0.00012251586304046214
step: 310, loss: 6.828539335401729e-05
step: 320, loss: 2.0186891561024822e-05
step: 330, loss: 3.633794767665677e-05
step: 340, loss: 0.00014479397214017808
step: 350, loss: 0.0001359164743917063
step: 360, loss: 0.0008567713666707277
step: 370, loss: 8.025446004467085e-05
step: 380, loss: 7.942911906866357e-05
step: 390, loss: 9.477381536271423e-05
step: 400, loss: 2.43443064391613e-05
step: 410, loss: 0.0004709537897724658
step: 420, loss: 0.005941905081272125
step: 430, loss: 0.004216314293444157
step: 440, loss: 0.0001228846376761794
step: 450, loss: 3.854314491036348e-05
step: 460, loss: 1.785881795512978e-05
step: 470, loss: 0.0004735949041787535
step: 480, loss: 0.00014676403952762485
step: 490, loss: 0.0011026135180145502
step: 500, loss: 6.935171404620633e-05
step: 510, loss: 0.00014025942073203623
step: 520, loss: 2.894820863730274e-05
step: 530, loss: 4.199921022518538e-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=0.9284369114877589, f1=0.9265325222274216,
best_f1=0.9222988505747127
```


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

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

step: 0, loss: 0.001372679602354765
step: 10, loss: 2.418760777800344e-05
step: 20, loss: 1.7486261640442535e-05
step: 30, loss: 0.0005396975320763886
step: 40, loss: 4.473160515772179e-05
step: 50, loss: 9.143666102318093e-05
step: 60, loss: 7.117149652913213e-05
step: 70, loss: 5.220070306677371e-05
step: 80, loss: 2.0525969375739805e-05
step: 90, loss: 0.0053299907594919205
step: 100, loss: 0.00011098036338808015
step: 110, loss: 0.00010036156891146675
step: 120, loss: 5.3816438594367355e-05
step: 130, loss: 2.0626519471989013e-05
step: 140, loss: 3.3492502552689984e-05
step: 150, loss: 0.0020242012105882168
step: 160, loss: 2.2250607798923738e-05
step: 170, loss: 0.00035261319135315716
step: 180, loss: 3.503131665638648e-05
step: 190, loss: 2.158362985937856e-05
step: 200, loss: 2.143491292372346e-05
step: 210, loss: 2.800549555104226e-05
step: 220, loss: 1.601482290425338e-05
step: 230, loss: 3.6290664866101e-05
step: 240, loss: 0.0002467202430125326
step: 250, loss: 2.0205601686029695e-05
step: 260, loss: 3.1814819521969184e-05
step: 270, loss: 1.616750523680821e-05
step: 280, loss: 2.2347408958012238e-05
step: 290, loss: 3.355910666869022e-05
step: 300, loss: 0.0002918803074862808
step: 310, loss: 0.0009047173080034554
step: 320, loss: 0.001981559442356229
step: 330, loss: 0.0005207728245295584
step: 340, loss: 2.634401425893884e-05
step: 350, loss: 3.144257789244875e-05
step: 360, loss: 2.948559085780289e-05
step: 370, loss: 3.16068617394194e-05
step: 380, loss: 3.62726605089847e-05
step: 390, loss: 0.00015168561367318034
step: 400, loss: 3.0036908356123604e-05
step: 410, loss: 0.00015940114099066705
step: 420, loss: 1.670027086220216e-05
step: 430, loss: 0.0013448743848130107
step: 440, loss: 2.5416617063456215e-05

```

step: 450, loss: 7.910388376330957e-05
step: 460, loss: 0.0006614992744289339
step: 470, loss: 1.3325223335414194e-05
step: 480, loss: 2.428340849291999e-05
step: 490, loss: 3.4791326470440254e-05
step: 500, loss: 1.3135238987160847e-05
step: 510, loss: 2.1247824406600557e-05
step: 520, loss: 0.00018942367751151323
step: 530, loss: 1.823107959353365e-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=0.9299303944315545, f1=0.9252767527675276,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002359656849876046
step: 10, loss: 2.1214887965470552e-05
step: 20, loss: 0.0007738119456917048
step: 30, loss: 0.00010745474719442427
step: 40, loss: 0.0005091029452160001
step: 50, loss: 0.00028053909773007035
step: 60, loss: 1.645434895181097e-05
step: 70, loss: 1.760919258231297e-05
step: 80, loss: 8.012853504624218e-05
step: 90, loss: 3.265397026552819e-05
step: 100, loss: 0.005030362866818905
step: 110, loss: 3.657710476545617e-05
step: 120, loss: 3.641576040536165e-05
step: 130, loss: 0.02732202783226967
step: 140, loss: 4.923817323287949e-05
step: 150, loss: 1.8205177184427157e-05
step: 160, loss: 5.3155537898419425e-05
step: 170, loss: 0.00015877022815402597
step: 180, loss: 2.370304719079286e-05
step: 190, loss: 0.0032022136729210615
step: 200, loss: 3.113873754045926e-05
step: 210, loss: 2.0578179828589782e-05
step: 220, loss: 2.136769035132602e-05
step: 230, loss: 0.0016770681831985712
step: 240, loss: 4.5125405449653044e-05
step: 250, loss: 2.6124660507775843e-05
step: 260, loss: 3.605367601267062e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 270, loss: 5.6708133342908695e-05
step: 280, loss: 4.7821322368690744e-05
step: 290, loss: 5.8124114730162546e-05

```

```

step: 300, loss: 2.225071693828795e-05
step: 310, loss: 0.00029828035621903837
step: 320, loss: 0.00012964929919689894
step: 330, loss: 5.704248178517446e-05
step: 340, loss: 3.50222981069237e-05
step: 350, loss: 0.00011240167077630758
step: 360, loss: 0.06337511539459229
step: 370, loss: 2.558783489803318e-05
step: 380, loss: 2.4813423806335777e-05
step: 390, loss: 0.004944520071148872
step: 400, loss: 5.191111995372921e-05
step: 410, loss: 1.8134125639335252e-05
step: 420, loss: 2.9051407182123512e-05
step: 430, loss: 6.616143218707293e-05
step: 440, loss: 8.440229430561885e-05
step: 450, loss: 7.559170626336709e-05
step: 460, loss: 2.599348772491794e-05
step: 470, loss: 1.828705171647016e-05
step: 480, loss: 2.2254038412938826e-05
step: 490, loss: 1.874533154477831e-05
step: 500, loss: 1.4975514204706997e-05
step: 510, loss: 0.01219907682389021
step: 520, loss: 4.823922427021898e-05
step: 530, loss: 2.7684996894095093e-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=0.9299905392620624, f1=0.9257518796992481,
best_f1=0.9222988505747127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.178856811951846e-05
step: 10, loss: 1.655498090258334e-05
step: 20, loss: 2.7081228836323135e-05
step: 30, loss: 0.00018950695812236518
step: 40, loss: 0.009091636165976524
step: 50, loss: 4.28300081694033e-05
step: 60, loss: 3.564063081284985e-05
step: 70, loss: 0.002760433591902256
step: 80, loss: 2.8909586035297252e-05
step: 90, loss: 2.9628974516526796e-05
step: 100, loss: 2.3509121092502028e-05
step: 110, loss: 5.68201721762307e-05
step: 120, loss: 0.0005630645900964737
step: 130, loss: 0.050705261528491974
step: 140, loss: 8.820874791126698e-05
step: 150, loss: 1.910296850837767e-05

```

```

step: 160, loss: 9.069030056707561e-05
step: 170, loss: 0.000134886140585877
step: 180, loss: 3.662397284642793e-05
step: 190, loss: 4.559089938993566e-05
step: 200, loss: 1.8171520423493348e-05
step: 210, loss: 0.0003102347836829722
step: 220, loss: 2.3658307327423245e-05
step: 230, loss: 3.305726932012476e-05
step: 240, loss: 4.398372766445391e-05
step: 250, loss: 2.311823845957406e-05
step: 260, loss: 1.5817422536201775e-05
step: 270, loss: 1.5213919141388033e-05
step: 280, loss: 3.4248583688167855e-05
step: 290, loss: 1.2978801350982394e-05
step: 300, loss: 1.6148931536008604e-05
step: 310, loss: 0.00010439592006150633
step: 320, loss: 1.7776772438082844e-05
step: 330, loss: 1.7828871932579204e-05
step: 340, loss: 1.3276821846375242e-05
step: 350, loss: 1.5470810467377305e-05
step: 360, loss: 0.0020314217545092106
step: 370, loss: 1.5146791156439576e-05
step: 380, loss: 3.820312849711627e-05
step: 390, loss: 4.577890285872854e-05
step: 400, loss: 3.617135007516481e-05
step: 410, loss: 2.7073750970885158e-05
step: 420, loss: 2.689815482881386e-05
step: 430, loss: 2.9517474104068242e-05
step: 440, loss: 1.9151428205077536e-05
step: 450, loss: 1.673365659371484e-05
step: 460, loss: 1.8804876162903383e-05
step: 470, loss: 0.000165477380505763
step: 480, loss: 1.2908030839753337e-05
step: 490, loss: 1.4681198990729172e-05
step: 500, loss: 4.272803198546171e-05
step: 510, loss: 2.33196078625042e-05
step: 520, loss: 1.5038816854939796e-05
step: 530, loss: 4.012647696072236e-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=0.9314553990610329, f1=0.9270346117867166,
best_f1=0.9222988505747127
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 [00:30, 185.84it/s]
load_f1 = 0.9282739472466451
real_f1 = 0.9265799256505577
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 177.47it/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: 0.559829592704773
step: 10, loss: 0.35776859521865845
step: 20, loss: 0.39042720198631287
step: 30, loss: 0.3186863362789154
step: 40, loss: 0.17897576093673706
step: 50, loss: 0.4299011528491974
step: 60, loss: 0.3272276222705841
step: 70, loss: 0.18039266765117645
step: 80, loss: 0.18885678052902222
step: 90, loss: 0.4533100426197052
step: 100, loss: 0.3786313831806183
step: 110, loss: 0.19130678474903107
step: 120, loss: 0.2685830295085907
step: 130, loss: 0.15352457761764526
step: 140, loss: 0.3449651002883911
step: 150, loss: 0.2473987638950348
step: 160, loss: 0.3724004626274109
step: 170, loss: 0.22690705955028534
step: 180, loss: 0.15363478660583496
step: 190, loss: 0.19710369408130646
step: 200, loss: 0.2925253212451935
step: 210, loss: 0.20270362496376038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6018348623853211, f1=0.6021897810218978,
best_f1=0.6021897810218978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12574860453605652
step: 10, loss: 0.17844849824905396
step: 20, loss: 0.17750893533229828
step: 30, loss: 0.21541379392147064
step: 40, loss: 0.19454790651798248

```

```

step: 50, loss: 0.19529353082180023
step: 60, loss: 0.4758639931678772
step: 70, loss: 0.1465493142604828
step: 80, loss: 0.19298091530799866
step: 90, loss: 0.09367386996746063
step: 100, loss: 0.028457511216402054
step: 110, loss: 0.13631656765937805
step: 120, loss: 0.1633710414171219
step: 130, loss: 0.05323651060461998
step: 140, loss: 0.1816863715648651
step: 150, loss: 0.25186797976493835
step: 160, loss: 0.19961333274841309
step: 170, loss: 0.10729032009840012
step: 180, loss: 0.18284305930137634
step: 190, loss: 0.16727280616760254
step: 200, loss: 0.14718946814537048
step: 210, loss: 0.1958560198545456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6183206106870229, f1=0.6435452793834296,
best_f1=0.6435452793834296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09317246824502945
step: 10, loss: 0.13484583795070648
step: 20, loss: 0.08088520914316177
step: 30, loss: 0.11232322454452515
step: 40, loss: 0.21535876393318176
step: 50, loss: 0.03610612452030182
step: 60, loss: 0.15268006920814514
step: 70, loss: 0.15341807901859283
step: 80, loss: 0.18252821266651154
step: 90, loss: 0.05264953523874283
step: 100, loss: 0.19984206557273865
step: 110, loss: 0.13932503759860992
step: 120, loss: 0.08967564254999161
step: 130, loss: 0.11559707671403885
step: 140, loss: 0.1335253268480301
step: 150, loss: 0.26776447892189026
step: 160, loss: 0.04518653452396393
step: 170, loss: 0.1609368622303009
step: 180, loss: 0.09510163217782974

```



```

step: 190, loss: 0.15319523215293884
step: 200, loss: 0.0805501714348793
step: 210, loss: 0.08823201060295105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6273764258555133, f1=0.6412825651302605,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11995422095060349
step: 10, loss: 0.059032220393419266
step: 20, loss: 0.07176954299211502
step: 30, loss: 0.06019839271903038
step: 40, loss: 0.024808313697576523
step: 50, loss: 0.18353614211082458
step: 60, loss: 0.2340761125087738
step: 70, loss: 0.23058989644050598
step: 80, loss: 0.09006053954362869
step: 90, loss: 0.05483076721429825
step: 100, loss: 0.19351772964000702
step: 110, loss: 0.13253122568130493
step: 120, loss: 0.1307401955127716
step: 130, loss: 0.12883317470550537
step: 140, loss: 0.08895077556371689
step: 150, loss: 0.023212725296616554
step: 160, loss: 0.050493981689214706
step: 170, loss: 0.07662392407655716
step: 180, loss: 0.26482880115509033
step: 190, loss: 0.07217232882976532
step: 200, loss: 0.15220418572425842
step: 210, loss: 0.12020494788885117
/usr/local/lib/python3.7/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=0.6213991769547326, f1=0.6279569892473118,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11448584496974945
step: 10, loss: 0.046234145760536194
step: 20, loss: 0.2947741150856018

```

```

step: 30, loss: 0.08354359865188599
step: 40, loss: 0.06284034997224808
step: 50, loss: 0.24151821434497833
step: 60, loss: 0.2293650209903717
step: 70, loss: 0.22479480504989624
step: 80, loss: 0.24923983216285706
step: 90, loss: 0.08424267172813416
step: 100, loss: 0.024092217907309532
step: 110, loss: 0.3609181046485901
step: 120, loss: 0.19254416227340698
step: 130, loss: 0.10329414159059525
step: 140, loss: 0.117304727435112
step: 150, loss: 0.2606677711009979
step: 160, loss: 0.10258123278617859
step: 170, loss: 0.10274742543697357
step: 180, loss: 0.1348002403974533
step: 190, loss: 0.022876542061567307
step: 200, loss: 0.05991535633802414
step: 210, loss: 0.00906207226216793
/usr/local/lib/python3.7/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=0.6203007518796992, f1=0.625, best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04331550374627113
step: 10, loss: 0.07968304306268692
step: 20, loss: 0.032599128782749176
step: 30, loss: 0.0024082453455775976
step: 40, loss: 0.060487568378448486
step: 50, loss: 0.021975858137011528
step: 60, loss: 0.15264420211315155
step: 70, loss: 0.007188237737864256
step: 80, loss: 0.10273823142051697
step: 90, loss: 0.195957213640213
step: 100, loss: 0.00803341530263424
step: 110, loss: 0.07057903707027435
step: 120, loss: 0.050091709941625595
step: 130, loss: 0.046031128615140915
step: 140, loss: 0.08132356405258179
step: 150, loss: 0.039233241230249405
step: 160, loss: 0.021033912897109985
step: 170, loss: 0.08800260722637177
step: 180, loss: 0.010479362681508064
step: 190, loss: 0.07149553298950195
step: 200, loss: 0.02541089989244938
step: 210, loss: 0.0052398014813661575

```

```

/usr/local/lib/python3.7/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=0.6243194192377496, f1=0.6497064579256361,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006910440046340227
step: 10, loss: 0.04854799807071686
step: 20, loss: 0.006619859021157026
step: 30, loss: 0.011738290078938007
step: 40, loss: 0.11927629262208939
step: 50, loss: 0.10439682006835938
step: 60, loss: 0.048819173127412796
step: 70, loss: 0.013168424367904663
step: 80, loss: 0.03828953579068184
step: 90, loss: 0.15870708227157593
step: 100, loss: 0.039869654923677444
step: 110, loss: 0.18083828687667847
step: 120, loss: 0.06951256096363068
step: 130, loss: 0.2034972906112671
step: 140, loss: 0.13993923366069794
step: 150, loss: 0.01416271273046732
step: 160, loss: 0.040799420326948166
step: 170, loss: 0.01577628217637539
step: 180, loss: 0.07724988460540771
step: 190, loss: 0.12268868088722229
step: 200, loss: 0.03923599794507027
step: 210, loss: 0.02939586341381073
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 7: dev_f1=0.6116838487972508, f1=0.59391771019678,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06745665520429611
step: 10, loss: 0.06572891771793365
step: 20, loss: 0.012342211790382862
step: 30, loss: 0.08709292858839035
step: 40, loss: 0.025871656835079193
step: 50, loss: 0.0251813605427742
step: 60, loss: 0.06988243013620377
step: 70, loss: 0.02601679600775242
step: 80, loss: 0.023583395406603813

```

```

step: 90, loss: 0.007538295350968838
step: 100, loss: 0.09433901309967041
step: 110, loss: 0.06031635031104088
step: 120, loss: 0.0036945920437574387
step: 130, loss: 0.003519246820360422
step: 140, loss: 0.08927781879901886
step: 150, loss: 0.0037123365327715874
step: 160, loss: 0.017537791281938553
step: 170, loss: 0.06770670413970947
step: 180, loss: 0.0931725725531578
step: 190, loss: 0.016220007091760635
step: 200, loss: 0.05031910911202431
step: 210, loss: 0.12452200800180435
/usr/local/lib/python3.7/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=0.6135922330097088, f1=0.6134969325153374,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00246094330213964
step: 10, loss: 0.10758835822343826
step: 20, loss: 0.033909041434526443
step: 30, loss: 0.047613371163606644
step: 40, loss: 0.19224314391613007
step: 50, loss: 0.08148784190416336
step: 60, loss: 0.14618459343910217
step: 70, loss: 0.00969864521175623
step: 80, loss: 0.06483925879001617
step: 90, loss: 0.004139646887779236
step: 100, loss: 0.001787970308214426
step: 110, loss: 0.016557883471250534
step: 120, loss: 0.017126860097050667
step: 130, loss: 0.030833708122372627
step: 140, loss: 0.029075928032398224
step: 150, loss: 0.3052505850791931
step: 160, loss: 0.008951952680945396
step: 170, loss: 0.13689586520195007
step: 180, loss: 0.03422942012548447
step: 190, loss: 0.0014097891980782151
step: 200, loss: 0.013234484940767288
step: 210, loss: 0.006467558443546295
/usr/local/lib/python3.7/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=0.6204238921001927, f1=0.6061855670103093,
best_f1=0.6412825651302605

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01427327562123537
step: 10, loss: 0.27994051575660706
step: 20, loss: 0.001233346643857658
step: 30, loss: 0.04699354246258736
step: 40, loss: 0.0007948959246277809
step: 50, loss: 0.003112548030912876
step: 60, loss: 0.02149033360183239
step: 70, loss: 0.06540756672620773
step: 80, loss: 0.012947739101946354
step: 90, loss: 0.03781115263700485
step: 100, loss: 0.04017501324415207
step: 110, loss: 0.010996662080287933
step: 120, loss: 0.01174743752926588
step: 130, loss: 0.027638860046863556
step: 140, loss: 0.0011542070424184203
step: 150, loss: 0.16462579369544983
step: 160, loss: 0.06449762731790543
step: 170, loss: 0.004935525357723236
step: 180, loss: 0.21298734843730927
step: 190, loss: 0.08709767460823059
step: 200, loss: 0.012541464529931545
step: 210, loss: 0.13217133283615112
/usr/local/lib/python3.7/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=0.5826771653543307, f1=0.6012526096033404,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02400214783847332
step: 10, loss: 0.029684890061616898
step: 20, loss: 0.014870557002723217
step: 30, loss: 0.0008163708262145519
step: 40, loss: 0.07255128771066666
step: 50, loss: 0.0030096168629825115
step: 60, loss: 0.10529438406229019
step: 70, loss: 0.0033130634110420942
step: 80, loss: 0.07199306041002274
step: 90, loss: 0.008342739194631577
step: 100, loss: 0.10553369671106339
step: 110, loss: 0.002644707215949893
step: 120, loss: 0.01621069200336933
step: 130, loss: 0.006892361678183079
step: 140, loss: 0.00349289714358747

```

```

step: 150, loss: 0.0045784576795995235
step: 160, loss: 0.0012221290962770581
step: 170, loss: 0.0036168701481074095
step: 180, loss: 0.004953187424689531
step: 190, loss: 0.0024082637391984463
step: 200, loss: 0.0018022630829364061
step: 210, loss: 0.0121039729565382
/usr/local/lib/python3.7/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=0.6035502958579881, f1=0.6157112526539279,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0049063158221542835
step: 10, loss: 0.002793485764414072
step: 20, loss: 0.01104117464274168
step: 30, loss: 0.04448234289884567
step: 40, loss: 0.18234257400035858
step: 50, loss: 0.007332105189561844
step: 60, loss: 0.02411622740328312
step: 70, loss: 0.003658879781141877
step: 80, loss: 0.001604176126420498
step: 90, loss: 0.023240922018885612
step: 100, loss: 0.007232555653899908
step: 110, loss: 0.019063841551542282
step: 120, loss: 0.010915701277554035
step: 130, loss: 0.005922381300479174
step: 140, loss: 0.00047436487511731684
step: 150, loss: 0.014637229964137077
step: 160, loss: 0.059326671063899994
step: 170, loss: 0.006898671854287386
step: 180, loss: 0.0029452317394316196
step: 190, loss: 0.004076004959642887
step: 200, loss: 0.012017332017421722
step: 210, loss: 0.021309897303581238
/usr/local/lib/python3.7/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=0.592734225621415, f1=0.6088709677419355,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09238404780626297
step: 10, loss: 0.017514409497380257
step: 20, loss: 0.04014328494668007

```

```

step: 30, loss: 0.100232794880867
step: 40, loss: 0.0030118864960968494
step: 50, loss: 0.18575164675712585
step: 60, loss: 0.004096794407814741
step: 70, loss: 0.014701404608786106
step: 80, loss: 0.013660695403814316
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.0014107715105637908
step: 100, loss: 0.0012931954115629196
step: 110, loss: 0.0030788076110184193
step: 120, loss: 0.005915038753300905
step: 130, loss: 0.032180674374103546
step: 140, loss: 0.01194145530462265
step: 150, loss: 0.013425626792013645
step: 160, loss: 0.010564970783889294
step: 170, loss: 0.0019470155239105225
step: 180, loss: 0.129952073097229
step: 190, loss: 0.013345077633857727
step: 200, loss: 0.0022687159944325686
step: 210, loss: 0.004499695263803005
/usr/local/lib/python3.7/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=0.5950095969289827, f1=0.601593625498008,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007945649558678269
step: 10, loss: 0.009926346130669117
step: 20, loss: 0.0064526633359491825
step: 30, loss: 0.003687534248456359
step: 40, loss: 0.0035970236640423536
step: 50, loss: 0.0034843445755541325
step: 60, loss: 0.0021747162099927664
step: 70, loss: 0.025172870606184006
step: 80, loss: 0.01103243138641119
step: 90, loss: 0.0207555890083313
step: 100, loss: 0.019083837047219276
step: 110, loss: 0.0014865281991660595
step: 120, loss: 0.020500605925917625
step: 130, loss: 0.0017621406586840749
step: 140, loss: 0.032581206411123276
step: 150, loss: 0.0093985041603446
step: 160, loss: 0.016708925366401672
step: 170, loss: 0.021064331755042076
step: 180, loss: 0.0016234348295256495
step: 190, loss: 0.0019423733465373516

```

```

step: 200, loss: 0.03774555027484894
step: 210, loss: 0.017470024526119232
/usr/local/lib/python3.7/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=0.6040515653775322, f1=0.607843137254902,
best_f1=0.6412825651302605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004411890637129545
step: 10, loss: 0.001025805831886828
step: 20, loss: 0.007354225032031536
step: 30, loss: 0.02948819287121296
step: 40, loss: 0.001547246240079403
step: 50, loss: 0.0012607157696038485
step: 60, loss: 0.28747981786727905
step: 70, loss: 0.0008451069006696343
step: 80, loss: 0.001054777530953288
step: 90, loss: 0.01243531983345747
step: 100, loss: 0.003116788575425744
step: 110, loss: 0.06164657324552536
step: 120, loss: 0.004041338339447975
step: 130, loss: 0.3406500220298767
step: 140, loss: 0.002521826885640621
step: 150, loss: 0.0009397291578352451
step: 160, loss: 0.015117661096155643
step: 170, loss: 0.004524680320173502
step: 180, loss: 0.0003568529209587723
step: 190, loss: 0.0033984980545938015
step: 200, loss: 0.01065011601895094
step: 210, loss: 0.007974245585501194
/usr/local/lib/python3.7/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=0.6003683241252302, f1=0.6062992125984252,
best_f1=0.6412825651302605
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, 261.13it/s]
load_f1 = 0.6290322580645161
real_f1 = 0.6285714285714287
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, 188.81it/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 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: 0.5747236013412476
step: 10, loss: 0.3638470470905304
step: 20, loss: 0.2946171164512634
step: 30, loss: 0.4313517212867737
step: 40, loss: 0.431461900472641
step: 50, loss: 0.2898389995098114
step: 60, loss: 0.26397329568862915
step: 70, loss: 0.2432250827550888
step: 80, loss: 0.2559962570667267
step: 90, loss: 0.29242852330207825
step: 100, loss: 0.3037935495376587
step: 110, loss: 0.5626217126846313
step: 120, loss: 0.14348027110099792
step: 130, loss: 0.13336396217346191
step: 140, loss: 0.07332278788089752
step: 150, loss: 0.15631060302257538
step: 160, loss: 0.14427506923675537
step: 170, loss: 0.22846178710460663
step: 180, loss: 0.10320916771888733
step: 190, loss: 0.2561503052711487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7055702917771883, f1=0.7112299465240641,
best_f1=0.7112299465240641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.295270174741745
step: 10, loss: 0.09230179339647293
step: 20, loss: 0.21448473632335663
step: 30, loss: 0.11265292763710022
step: 40, loss: 0.1952105015516281
step: 50, loss: 0.02495179884135723
step: 60, loss: 0.15653595328330994
step: 70, loss: 0.2365403026342392
step: 80, loss: 0.14229846000671387
step: 90, loss: 0.15057721734046936
step: 100, loss: 0.015796279534697533
step: 110, loss: 0.1645171046257019
step: 120, loss: 0.31518158316612244

```

```

step: 130, loss: 0.04722382128238678
step: 140, loss: 0.06716885417699814
step: 150, loss: 0.11182794719934464
step: 160, loss: 0.13478714227676392
step: 170, loss: 0.09588626772165298
step: 180, loss: 0.10214047878980637
step: 190, loss: 0.025740914046764374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7589498806682579, f1=0.7310344827586207,
best_f1=0.7310344827586207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03886287286877632
step: 10, loss: 0.165261909365654
step: 20, loss: 0.12429823726415634
step: 30, loss: 0.02047938108444214
step: 40, loss: 0.031813520938158035
step: 50, loss: 0.08257211744785309
step: 60, loss: 0.06522100418806076
step: 70, loss: 0.14192283153533936
step: 80, loss: 0.07181935012340546
step: 90, loss: 0.10299234837293625
step: 100, loss: 0.10679931938648224
step: 110, loss: 0.022499749436974525
step: 120, loss: 0.10645530372858047
step: 130, loss: 0.022092550992965698
step: 140, loss: 0.016562819480895996
step: 150, loss: 0.13293829560279846
step: 160, loss: 0.032913465052843094
step: 170, loss: 0.13016541302204132
step: 180, loss: 0.061252661049366
step: 190, loss: 0.05525524169206619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7603305785123966, f1=0.7645429362880887,
best_f1=0.7645429362880887

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03176116570830345
step: 10, loss: 0.11215424537658691
step: 20, loss: 0.008654404431581497
step: 30, loss: 0.011429822072386742
step: 40, loss: 0.022609923034906387
step: 50, loss: 0.06457199901342392
step: 60, loss: 0.014971532858908176
step: 70, loss: 0.23135961592197418
step: 80, loss: 0.15465039014816284
step: 90, loss: 0.016582228243350983
step: 100, loss: 0.0705464705824852
step: 110, loss: 0.00744925532490015
step: 120, loss: 0.15323418378829956
step: 130, loss: 0.1531699001789093
step: 140, loss: 0.015389365144073963
step: 150, loss: 0.037868086248636246
step: 160, loss: 0.056008562445640564
step: 170, loss: 0.08883054554462433
step: 180, loss: 0.01867492124438286
step: 190, loss: 0.28805580735206604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7765957446808511, f1=0.7708894878706198,
best_f1=0.7708894878706198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11971718072891235
step: 10, loss: 0.029741816222667694
step: 20, loss: 0.05840824916958809
step: 30, loss: 0.007809980306774378
step: 40, loss: 0.053557198494672775
step: 50, loss: 0.11741618067026138
step: 60, loss: 0.13429324328899384
step: 70, loss: 0.025861380621790886
step: 80, loss: 0.05446785315871239
step: 90, loss: 0.008069389499723911
step: 100, loss: 0.0036188322119414806
step: 110, loss: 0.013332745060324669
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```

```

step: 120, loss: 0.14618515968322754
step: 130, loss: 0.041115809231996536
step: 140, loss: 0.018154626712203026
step: 150, loss: 0.007402552291750908
step: 160, loss: 0.08700843900442123
step: 170, loss: 0.013060525991022587
step: 180, loss: 0.01994853839278221
step: 190, loss: 0.11875659227371216
/usr/local/lib/python3.7/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=0.7456647398843931, f1=0.7603305785123966,
best_f1=0.7708894878706198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009288841858506203
step: 10, loss: 0.15127800405025482
step: 20, loss: 0.019703608006238937
step: 30, loss: 0.0009264039108529687
step: 40, loss: 0.11951709538698196
step: 50, loss: 0.15691593289375305
step: 60, loss: 0.049912914633750916
step: 70, loss: 0.05108461156487465
step: 80, loss: 0.011548801325261593
step: 90, loss: 0.006949618458747864
step: 100, loss: 0.0012812769273295999
step: 110, loss: 0.018597444519400597
step: 120, loss: 0.007441353052854538
step: 130, loss: 0.03729420527815819
step: 140, loss: 0.1789143681526184
step: 150, loss: 0.004625928122550249
step: 160, loss: 0.04423690587282181
step: 170, loss: 0.04422051087021828
step: 180, loss: 0.018787488341331482
step: 190, loss: 0.2225644290447235
/usr/local/lib/python3.7/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=0.7590027700831025, f1=0.765498652291105,
best_f1=0.7708894878706198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024834826588630676
step: 10, loss: 0.018124382942914963
step: 20, loss: 0.11240343004465103
step: 30, loss: 0.007269229274243116

```

```

step: 40, loss: 0.10424968600273132
step: 50, loss: 0.05602165311574936
step: 60, loss: 0.01984022930264473
step: 70, loss: 0.004588773939758539
step: 80, loss: 0.06548745185136795
step: 90, loss: 0.014920404180884361
step: 100, loss: 0.0018296337220817804
step: 110, loss: 0.021954642608761787
step: 120, loss: 0.0019309065537527204
step: 130, loss: 0.049784742295742035
step: 140, loss: 0.083681620657444
step: 150, loss: 0.01589987613260746
step: 160, loss: 0.0318194255232811
step: 170, loss: 0.0017467293655499816
step: 180, loss: 0.006374107673764229
step: 190, loss: 0.0030625779181718826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.783289817232376, f1=0.7864583333333334,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05195891857147217
step: 10, loss: 0.05226174741983414
step: 20, loss: 0.0012838663533329964
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0018832386704161763
step: 40, loss: 0.0009554364951327443
step: 50, loss: 0.00393577991053462
step: 60, loss: 0.001337807858362794
step: 70, loss: 0.004316406790167093
step: 80, loss: 0.00043436125270090997
step: 90, loss: 0.004922202322632074
step: 100, loss: 0.037031255662441254
step: 110, loss: 0.0012711520539596677
step: 120, loss: 0.009230081923305988
step: 130, loss: 0.04704660177230835
step: 140, loss: 0.048279065638780594
step: 150, loss: 0.08298180997371674
step: 160, loss: 0.005188571754842997
step: 170, loss: 0.003530934453010559
step: 180, loss: 0.03257189691066742

```

```

step: 190, loss: 0.03970470279455185
/usr/local/lib/python3.7/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=0.7450980392156863, f1=0.7814207650273224,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004016768652945757
step: 10, loss: 0.053457748144865036
step: 20, loss: 0.0029473016038537025
step: 30, loss: 0.004308617208153009
step: 40, loss: 0.012330293655395508
step: 50, loss: 0.00314158177934587
step: 60, loss: 0.0009170083794742823
step: 70, loss: 0.0009093850967474282
step: 80, loss: 0.0038862922228872776
step: 90, loss: 0.047834884375333786
step: 100, loss: 0.0035457261838018894
step: 110, loss: 0.0014239661395549774
step: 120, loss: 0.015846529975533485
step: 130, loss: 0.0025319410488009453
step: 140, loss: 0.01223894115537405
step: 150, loss: 0.06211300194263458
step: 160, loss: 0.0006073106196708977
step: 170, loss: 0.003031275002285838
step: 180, loss: 0.0031873853877186775
step: 190, loss: 0.002588848117738962
/usr/local/lib/python3.7/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=0.7357512953367874, f1=0.7633587786259541,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006958133541047573
step: 10, loss: 0.018991561606526375
step: 20, loss: 0.038469646126031876
step: 30, loss: 0.0006795639055781066
step: 40, loss: 0.030146148055791855
step: 50, loss: 0.0009015586110763252
step: 60, loss: 0.004854666534811258
step: 70, loss: 0.0018037321278825402
step: 80, loss: 0.0008089960319921374
step: 90, loss: 0.00399294076487422
step: 100, loss: 0.062265701591968536

```



```

step: 110, loss: 0.0023801412899047136
step: 120, loss: 0.0357554629445076
step: 130, loss: 0.0018132735276594758
step: 140, loss: 0.0007521822117269039
step: 150, loss: 0.0020835495088249445
step: 160, loss: 0.004574147053062916
step: 170, loss: 0.001317595480941236
step: 180, loss: 0.0005500927800312638
step: 190, loss: 0.01732248067855835
/usr/local/lib/python3.7/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=0.7566137566137565, f1=0.783289817232376,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003821968275588006
step: 10, loss: 0.0019804849289357662
step: 20, loss: 0.03550595045089722
step: 30, loss: 0.0260370634496212
step: 40, loss: 0.0034650007728487253
step: 50, loss: 0.0033667266834527254
step: 60, loss: 0.00043802594882436097
step: 70, loss: 0.004694852977991104
step: 80, loss: 0.0008801307994872332
step: 90, loss: 0.002265841932967305
step: 100, loss: 0.0003444919129833579
step: 110, loss: 0.0013234724756330252
step: 120, loss: 0.0026577834505587816
step: 130, loss: 0.005270994268357754
step: 140, loss: 0.0004575580242089927
step: 150, loss: 0.0008810079889371991
step: 160, loss: 0.004362972918897867
step: 170, loss: 0.0005695278523489833
step: 180, loss: 0.0010966291883960366
step: 190, loss: 0.001476907404139638
/usr/local/lib/python3.7/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=0.7377049180327869, f1=0.7828418230563003,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039927856414578855
step: 10, loss: 0.015712466090917587
step: 20, loss: 0.0030201047193259

```

```

step: 30, loss: 0.0008365672547370195
step: 40, loss: 0.0017339398618787527
step: 50, loss: 0.03452081233263016
step: 60, loss: 0.0007609252352267504
step: 70, loss: 0.0013077151961624622
step: 80, loss: 0.0010516950860619545
step: 90, loss: 0.0009636275353841484
step: 100, loss: 0.0015973594272509217
step: 110, loss: 0.0022747728507965803
step: 120, loss: 0.03926264867186546
step: 130, loss: 0.006297870073467493
step: 140, loss: 0.0012183568906039
step: 150, loss: 0.0005265043000690639
step: 160, loss: 0.00022966426331549883
step: 170, loss: 0.0006011102814227343
step: 180, loss: 0.00021377943630795926
step: 190, loss: 0.0008079196559265256
/usr/local/lib/python3.7/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=0.7553191489361701, f1=0.7862796833773087,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00101335602812469
step: 10, loss: 0.0012214835733175278
step: 20, loss: 0.00044338623411022127
step: 30, loss: 0.0008343024528585374
step: 40, loss: 0.0003450032672844827
step: 50, loss: 0.027049703523516655
step: 60, loss: 0.0004750251828227192
step: 70, loss: 0.004563787952065468
step: 80, loss: 0.0008668092777952552
step: 90, loss: 0.0006108786328695714
step: 100, loss: 0.0003043382312171161
step: 110, loss: 0.16960668563842773
step: 120, loss: 0.007802753243595362
step: 130, loss: 0.000542774680070579
step: 140, loss: 0.0004172123153693974
step: 150, loss: 0.0010260781273245811
step: 160, loss: 0.0010999090736731887
step: 170, loss: 0.00340235885232687
step: 180, loss: 0.0005073210340924561
step: 190, loss: 0.009664095006883144
/usr/local/lib/python3.7/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=0.7407407407407407, f1=0.7835616438356164,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02005401812493801
step: 10, loss: 0.0008349038544110954
step: 20, loss: 0.0002009925083257258
step: 30, loss: 0.0003266129642724991
step: 40, loss: 0.0008473624475300312
step: 50, loss: 0.0008258640882559121
step: 60, loss: 0.0007931379368528724
step: 70, loss: 0.0006780883413739502
step: 80, loss: 0.0012964335037395358
step: 90, loss: 0.11589497327804565
step: 100, loss: 0.03202691674232483
step: 110, loss: 0.0023480383679270744
step: 120, loss: 0.0012164578074589372
step: 130, loss: 0.0010739531135186553
step: 140, loss: 0.0008553519146516919
step: 150, loss: 0.0003947324585169554
step: 160, loss: 0.0016354481922462583
step: 170, loss: 0.006974177900701761
step: 180, loss: 0.00349019980058074
step: 190, loss: 0.0008975538075901568
/usr/local/lib/python3.7/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=0.7409638554216867, f1=0.7610619469026548,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004958240315318108
step: 10, loss: 0.00043410234502516687
step: 20, loss: 0.0009233352611772716
step: 30, loss: 0.0007016299059614539
step: 40, loss: 0.0015881785657256842
step: 50, loss: 0.003964090254157782
step: 60, loss: 0.006848291028290987
step: 70, loss: 0.01478884182870388
step: 80, loss: 0.0016825946513563395
step: 90, loss: 0.003822986502200365
step: 100, loss: 0.00033926713513210416
step: 110, loss: 0.0016910646809265018
step: 120, loss: 0.0015290766023099422
step: 130, loss: 0.0005379270878620446
step: 140, loss: 0.004306087736040354

```

```

step: 150, loss: 0.0004184666904620826
step: 160, loss: 0.00046936215949244797
step: 170, loss: 0.0005404283874668181
step: 180, loss: 0.004987275693565607
step: 190, loss: 0.0008709009271115065
/usr/local/lib/python3.7/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=0.7388888888888889, f1=0.7780821917808218,
best_f1=0.7864583333333334
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, 164.93it/s]
load_f1 = 0.5522388059701493
real_f1 = 0.5253623188405796
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, 186.13it/s]

```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

```

After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6160624623298645
step: 10, loss: 0.3578070104122162
step: 20, loss: 0.3031207323074341
step: 30, loss: 0.40407049655914307
step: 40, loss: 0.2734251320362091
step: 50, loss: 0.23552265763282776
step: 60, loss: 0.32287082076072693
step: 70, loss: 0.37386763095855713
step: 80, loss: 0.3643961250782013
step: 90, loss: 0.2786964178085327
step: 100, loss: 0.17185182869434357
step: 110, loss: 0.2549002170562744
step: 120, loss: 0.23130127787590027
step: 130, loss: 0.07773682475090027
step: 140, loss: 0.20865677297115326
step: 150, loss: 0.32827886939048767
step: 160, loss: 0.18956556916236877
step: 170, loss: 0.1797092705965042
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6635730858468678, f1=0.6823529411764706,
best_f1=0.6823529411764706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1147431880235672
step: 10, loss: 0.17865954339504242
step: 20, loss: 0.0471271313726902
step: 30, loss: 0.15875324606895447
step: 40, loss: 0.14253467321395874
step: 50, loss: 0.218867227435112
step: 60, loss: 0.07670901715755463
step: 70, loss: 0.2089819759130478
step: 80, loss: 0.11561757326126099
step: 90, loss: 0.13069388270378113
step: 100, loss: 0.2010175883769989
step: 110, loss: 0.07862098515033722
step: 120, loss: 0.08925696462392807
step: 130, loss: 0.10106905549764633
step: 140, loss: 0.2541296184062958
step: 150, loss: 0.20542685687541962
step: 160, loss: 0.2103762924671173
step: 170, loss: 0.08257003873586655
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7563451776649747, f1=0.784503631961259,
best_f1=0.784503631961259
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14385943114757538
step: 10, loss: 0.03141995146870613
step: 20, loss: 0.038731470704078674
step: 30, loss: 0.15978626906871796
step: 40, loss: 0.10911650955677032
step: 50, loss: 0.14847300946712494
step: 60, loss: 0.14721040427684784
step: 70, loss: 0.0734223797917366
step: 80, loss: 0.04688633233308792
step: 90, loss: 0.11484850943088531
step: 100, loss: 0.016970839351415634
step: 110, loss: 0.09403610229492188

```

```

step: 120, loss: 0.13487356901168823
step: 130, loss: 0.10617193579673767
step: 140, loss: 0.11184470355510712
step: 150, loss: 0.030839012935757637
step: 160, loss: 0.03724424168467522
step: 170, loss: 0.10294406116008759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.772378516624041, f1=0.7884615384615384,
best_f1=0.7884615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02860618382692337
step: 10, loss: 0.025852227583527565
step: 20, loss: 0.0023072666954249144
step: 30, loss: 0.10964122414588928
step: 40, loss: 0.006808501202613115
step: 50, loss: 0.025144772604107857
step: 60, loss: 0.12931177020072937
step: 70, loss: 0.0035351132974028587
step: 80, loss: 0.017634117975831032
step: 90, loss: 0.06205407530069351
step: 100, loss: 0.1772691160440445
step: 110, loss: 0.06961572170257568
step: 120, loss: 0.021637756377458572
step: 130, loss: 0.027258306741714478
step: 140, loss: 0.014280815608799458
step: 150, loss: 0.08033650368452072
step: 160, loss: 0.19547638297080994
step: 170, loss: 0.05943191424012184
/usr/local/lib/python3.7/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=0.7609756097560975, f1=0.7895981087470448,
best_f1=0.7884615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025256626307964325
step: 10, loss: 0.0074385106563568115
step: 20, loss: 0.038938429206609726
step: 30, loss: 0.007590157445520163

```



```

step: 40, loss: 0.026013024151325226
step: 50, loss: 0.07185228914022446
step: 60, loss: 0.015393208712339401
step: 70, loss: 0.13099992275238037
step: 80, loss: 0.08710107952356339
step: 90, loss: 0.0809776559472084
step: 100, loss: 0.027644561603665352
step: 110, loss: 0.03333596512675285
step: 120, loss: 0.008177747018635273
step: 130, loss: 0.050626907497644424
step: 140, loss: 0.05312949791550636
step: 150, loss: 0.19108740985393524
step: 160, loss: 0.09564556926488876
step: 170, loss: 0.062960185110569
/usr/local/lib/python3.7/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=0.7692307692307693, f1=0.7769423558897243,
best_f1=0.7884615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027413146570324898
step: 10, loss: 0.007073075510561466
step: 20, loss: 0.012950733304023743
step: 30, loss: 0.007242200430482626
step: 40, loss: 0.017294298857450485
step: 50, loss: 0.11034324765205383
step: 60, loss: 0.0037846537306904793
step: 70, loss: 0.11395540833473206
step: 80, loss: 0.09265533834695816
step: 90, loss: 0.03480174019932747
step: 100, loss: 0.012932802550494671
step: 110, loss: 0.0015487887430936098
step: 120, loss: 0.007301964797079563
step: 130, loss: 0.01617995835840702
step: 140, loss: 0.01805436611175537
step: 150, loss: 0.05511849746108055
step: 160, loss: 0.1884027123451233
step: 170, loss: 0.006385842338204384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7837150127226464, f1=0.7892156862745098,

```

```

best_f1=0.7892156862745098
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00823309924453497
step: 10, loss: 0.041872888803482056
step: 20, loss: 0.002709451597183943
step: 30, loss: 0.012005547061562538
step: 40, loss: 0.018291160464286804
step: 50, loss: 0.009601768106222153
step: 60, loss: 0.0021386633161455393
step: 70, loss: 0.18515728414058685
step: 80, loss: 0.008033890277147293
step: 90, loss: 0.00039678646135143936
step: 100, loss: 0.002632558811455965
step: 110, loss: 0.0029281540773808956
step: 120, loss: 0.0009828683687373996
step: 130, loss: 0.14113228023052216
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.002929759444668889
step: 150, loss: 0.0023013290483504534
step: 160, loss: 0.005755018908530474
step: 170, loss: 0.04228533059358597
/usr/local/lib/python3.7/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=0.7535545023696684, f1=0.7820224719101124,
best_f1=0.7892156862745098
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.054257072508335114
step: 10, loss: 0.004721847362816334
step: 20, loss: 0.000993782072328031
step: 30, loss: 0.011911991983652115
step: 40, loss: 0.0006604493246413767
step: 50, loss: 0.007002763915807009
step: 60, loss: 0.000902554311323911
step: 70, loss: 0.0021249365527182817
step: 80, loss: 0.01895422674715519
step: 90, loss: 0.002239004010334611
step: 100, loss: 0.0404781736433506
step: 110, loss: 0.07070623338222504
step: 120, loss: 0.0009899426950141788
step: 130, loss: 0.0022594607435166836
step: 140, loss: 0.0010425494983792305
step: 150, loss: 0.057124823331832886
step: 160, loss: 0.021468980237841606

```

```

step: 170, loss: 0.0010089146671816707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7939698492462312, f1=0.7924528301886793,
best_f1=0.7924528301886793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006422956357710063
step: 10, loss: 0.017525367438793182
step: 20, loss: 0.013099067844450474
step: 30, loss: 0.0094574224203825
step: 40, loss: 0.005321857053786516
step: 50, loss: 0.0004915713798254728
step: 60, loss: 0.044678572565317154
step: 70, loss: 0.0010431779082864523
step: 80, loss: 0.000452656444394961
step: 90, loss: 0.01230384036898613
step: 100, loss: 0.00665422435849905
step: 110, loss: 0.001301756128668785
step: 120, loss: 0.01966683752834797
step: 130, loss: 0.13909536600112915
step: 140, loss: 0.0009327511070296168
step: 150, loss: 0.001398555003106594
step: 160, loss: 0.030807800590991974
step: 170, loss: 0.014537235721945763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7849462365591398, f1=0.7806122448979592,
best_f1=0.7924528301886793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020426468923687935
step: 10, loss: 0.05994807556271553
step: 20, loss: 0.0861298143863678
step: 30, loss: 0.003493951866403222
step: 40, loss: 0.0005925192963331938
step: 50, loss: 0.08383143693208694
step: 60, loss: 0.0018264725804328918
step: 70, loss: 0.006449682638049126
step: 80, loss: 0.016379227861762047

```

```

step: 90, loss: 0.02923404611647129
step: 100, loss: 0.0010468121618032455
step: 110, loss: 0.02825297974050045
step: 120, loss: 0.010149788111448288
step: 130, loss: 0.0006529054371640086
step: 140, loss: 0.032737959176301956
step: 150, loss: 0.015462280251085758
step: 160, loss: 0.10302317142486572
step: 170, loss: 0.0004631828924175352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7969151670951157, f1=0.7680798004987531,
best_f1=0.7680798004987531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03153115510940552
step: 10, loss: 0.08350678533315659
step: 20, loss: 0.00024414859944954515
step: 30, loss: 0.0002358444908168167
step: 40, loss: 0.002970406087115407
step: 50, loss: 0.011629598215222359
step: 60, loss: 0.013318079523742199
step: 70, loss: 0.0013037562603130937
step: 80, loss: 0.003489808179438114
step: 90, loss: 0.00028080071206204593
step: 100, loss: 0.0007029577973298728
step: 110, loss: 0.047434091567993164
step: 120, loss: 0.000543037080205977
step: 130, loss: 0.00013988149294164032
step: 140, loss: 0.01843039132654667
step: 150, loss: 0.054552506655454636
step: 160, loss: 0.011115096509456635
step: 170, loss: 0.0003409550408832729
/usr/local/lib/python3.7/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=0.7931034482758621, f1=0.7923627684964201,
best_f1=0.7680798004987531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03343457728624344

```

```

step: 10, loss: 0.0005043488345108926
step: 20, loss: 0.0015275615733116865
step: 30, loss: 0.0032992339693009853
step: 40, loss: 0.00020959859830327332
step: 50, loss: 0.002728637307882309
step: 60, loss: 0.00040208676364272833
step: 70, loss: 0.04172198846936226
step: 80, loss: 0.0001676696410868317
step: 90, loss: 0.001133674755692482
step: 100, loss: 0.0008063169079832733
step: 110, loss: 0.0005821168306283653
step: 120, loss: 0.028084512799978256
step: 130, loss: 0.004327119793742895
step: 140, loss: 0.0027129563968628645
step: 150, loss: 0.0032099010422825813
step: 160, loss: 0.0004068036796525121
step: 170, loss: 0.0016046181553974748
/usr/local/lib/python3.7/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=0.7883597883597883, f1=0.7774936061381074,
best_f1=0.7680798004987531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01176624745130539
step: 10, loss: 0.0005817834171466529
step: 20, loss: 0.0004980841185897589
step: 30, loss: 8.16402753116563e-05
step: 40, loss: 0.00023512494226451963
step: 50, loss: 0.000334053736878559
step: 60, loss: 0.0010646425653249025
step: 70, loss: 0.012457323260605335
step: 80, loss: 0.002598751336336136
step: 90, loss: 0.00012721741222776473
step: 100, loss: 0.00011289468966424465
step: 110, loss: 6.926029891474172e-05
step: 120, loss: 0.023852545768022537
step: 130, loss: 0.00031276908703148365
step: 140, loss: 0.00021103907783981413
step: 150, loss: 0.0006746590952388942
step: 160, loss: 0.0006202015792950988
step: 170, loss: 0.0003393988881725818
/usr/local/lib/python3.7/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=0.7801047120418847, f1=0.7877237851662404,
best_f1=0.7680798004987531

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018694688333198428
step: 10, loss: 0.00010724979074439034
step: 20, loss: 0.0010057686595246196
step: 30, loss: 0.0029675636906176805
step: 40, loss: 0.0008308644173666835
step: 50, loss: 0.0001815819850889966
step: 60, loss: 0.00010043365182355046
step: 70, loss: 0.00031637787469662726
step: 80, loss: 0.0008953088545240462
step: 90, loss: 0.009594021365046501
step: 100, loss: 0.015772048383951187
step: 110, loss: 0.00015931023517623544
step: 120, loss: 0.03544095531105995
step: 130, loss: 0.0010173969203606248
step: 140, loss: 0.00012654894089791924
step: 150, loss: 0.0003564849030226469
step: 160, loss: 0.00014504048158414662
step: 170, loss: 0.0014452841132879257
/usr/local/lib/python3.7/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=0.782608695652174, f1=0.7836538461538461,
best_f1=0.7680798004987531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002919707912951708
step: 10, loss: 0.0024327924475073814
step: 20, loss: 0.013726986944675446
step: 30, loss: 0.0002483545686118305
step: 40, loss: 0.0003644108073785901
step: 50, loss: 0.0010198104428127408
step: 60, loss: 0.009680924005806446
step: 70, loss: 0.00011388528218958527
step: 80, loss: 0.005458802450448275
step: 90, loss: 0.0008848047000356019
step: 100, loss: 0.0048545957542955875
step: 110, loss: 0.00011845019616885111
step: 120, loss: 0.011036824434995651
step: 130, loss: 0.00014448566071223468
step: 140, loss: 0.00018600614566821605
step: 150, loss: 0.0001583127595949918
step: 160, loss: 0.00016376070561818779
step: 170, loss: 0.00019657188386190683
/usr/local/lib/python3.7/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=0.782608695652174, f1=0.7853658536585365,
best_f1=0.7680798004987531
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:08, 230.33it/s]
load_f1 = 0.5894736842105264
real_f1 = 0.5838509316770186
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, 186.56it/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, 419kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 3.64MB/s]

Downloading: 100% 440M/440M [00:12<00:00, 35.9MB/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: 0.6235719919204712
step: 10, loss: 0.6312541365623474
step: 20, loss: 0.44503018260002136
step: 30, loss: 0.1743483543395996
step: 40, loss: 0.3155535161495209
step: 50, loss: 0.03272319585084915
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.03556232154369354
step: 70, loss: 0.07774319499731064
step: 80, loss: 0.033296555280685425
step: 90, loss: 0.1401262879371643
step: 100, loss: 0.008157103322446346
step: 110, loss: 0.2631863057613373
step: 120, loss: 0.01205198373645544
step: 130, loss: 0.06549979001283646
step: 140, loss: 0.07244355976581573
step: 150, loss: 0.022079085931181908
step: 160, loss: 0.008636778220534325
step: 170, loss: 0.14245842397212982
step: 180, loss: 0.004977049306035042
step: 190, loss: 0.022717732936143875

```

```

step: 200, loss: 0.14678341150283813
step: 210, loss: 0.02345052920281887
step: 220, loss: 0.03050711192190647
step: 230, loss: 0.09031583368778229
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9623893805309734, f1=0.9574944071588367,
best_f1=0.9574944071588367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013091005384922028
step: 10, loss: 0.12677666544914246
step: 20, loss: 0.26159656047821045
step: 30, loss: 0.21529702842235565
step: 40, loss: 0.04898177459836006
step: 50, loss: 0.010202880948781967
step: 60, loss: 0.04364316910505295
step: 70, loss: 0.032218411564826965
step: 80, loss: 0.011400094255805016
step: 90, loss: 0.027033142745494843
step: 100, loss: 0.12784825265407562
step: 110, loss: 0.11775511503219604
step: 120, loss: 0.010033784434199333
step: 130, loss: 0.026946881785988808
step: 140, loss: 0.013425392098724842
step: 150, loss: 0.015337568707764149
step: 160, loss: 0.023819509893655777
step: 170, loss: 0.0027103303000330925
step: 180, loss: 0.14464601874351501
step: 190, loss: 0.010040361434221268
step: 200, loss: 0.002524025971069932
step: 210, loss: 0.0013327037449926138
step: 220, loss: 0.037308190017938614
step: 230, loss: 0.05776024982333183
/usr/local/lib/python3.7/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=0.9592760180995475, f1=0.9511918274687856,
best_f1=0.9574944071588367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.013250243850052357
step: 10, loss: 0.004036106634885073
step: 20, loss: 0.061369236558675766
step: 30, loss: 0.07392865419387817
step: 40, loss: 0.12276484817266464
step: 50, loss: 0.02685686945915222
step: 60, loss: 0.005020324606448412
step: 70, loss: 0.012022594921290874
step: 80, loss: 0.002182349795475602
step: 90, loss: 0.16025806963443756
step: 100, loss: 0.0006016682600602508
step: 110, loss: 0.0006202809745445848
step: 120, loss: 0.030162647366523743
step: 130, loss: 0.0005714509170502424
step: 140, loss: 0.00392556469887495
step: 150, loss: 0.0033767998684197664
step: 160, loss: 0.0338832288980484
step: 170, loss: 0.08676934242248535
step: 180, loss: 0.0060205888003110886
step: 190, loss: 0.028023183345794678
step: 200, loss: 0.10408516973257065
step: 210, loss: 0.19926415383815765
step: 220, loss: 0.002148954663425684
step: 230, loss: 0.005834612064063549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9631284916201117, f1=0.9709821428571428,
best_f1=0.9709821428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017420463263988495
step: 10, loss: 0.0016254675574600697
step: 20, loss: 0.0021867582108825445
step: 30, loss: 0.0015318489167839289
step: 40, loss: 0.0027749286964535713
step: 50, loss: 0.0013703955337405205
step: 60, loss: 0.007128322031348944
step: 70, loss: 0.008927100338041782
step: 80, loss: 0.0004905008245259523
step: 90, loss: 0.0009122289484366775
step: 100, loss: 0.007226633373647928
step: 110, loss: 0.006021332461386919

```

```

step: 120, loss: 0.01563136838376522
step: 130, loss: 0.004285208415240049
step: 140, loss: 0.0024543567560613155
step: 150, loss: 0.1412256956100464
step: 160, loss: 0.06644104421138763
step: 170, loss: 0.033130258321762085
step: 180, loss: 0.004881853703409433
step: 190, loss: 0.008014014922082424
step: 200, loss: 0.0033086270559579134
step: 210, loss: 0.17250193655490875
step: 220, loss: 0.0005824373802170157
step: 230, loss: 0.028264909982681274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9668874172185431, f1=0.9656699889258028,
best_f1=0.9656699889258028
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019204160198569298
step: 10, loss: 0.011113934218883514
step: 20, loss: 0.13057874143123627
step: 30, loss: 0.0005883671110495925
step: 40, loss: 0.0014754958683624864
step: 50, loss: 0.012627651914954185
step: 60, loss: 0.007281249389052391
step: 70, loss: 0.0009360458352603018
step: 80, loss: 0.004042844753712416
step: 90, loss: 0.0010451892158016562
step: 100, loss: 0.00029603755683638155
step: 110, loss: 0.0003046469937544316
step: 120, loss: 0.00020256711286492646
step: 130, loss: 0.002675189170986414
step: 140, loss: 0.001932028797455132
step: 150, loss: 0.002173067070543766
step: 160, loss: 0.0020535450894385576
step: 170, loss: 0.0007931040017865598
step: 180, loss: 0.0019240269903093576
step: 190, loss: 0.1394355297088623
step: 200, loss: 0.00202564406208694
step: 210, loss: 0.021097412332892418
step: 220, loss: 0.0038689810317009687
step: 230, loss: 0.0007448342512361705

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.972129319955407, 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: 0.00432361988350749
step: 10, loss: 0.013641505502164364
step: 20, loss: 0.007067626807838678
step: 30, loss: 0.0002451637701597065
step: 40, loss: 0.0007077838527038693
step: 50, loss: 0.000759169168304652
step: 60, loss: 0.0002947907196357846
step: 70, loss: 0.001224396051838994
step: 80, loss: 0.001199872582219541
step: 90, loss: 0.005420915316790342
step: 100, loss: 0.022458994761109352
step: 110, loss: 0.0008206773782148957
step: 120, loss: 0.13915029168128967
step: 130, loss: 0.0003514142881613225
step: 140, loss: 0.003247681772336364
step: 150, loss: 0.1354210376739502
step: 160, loss: 0.0014391348231583834
step: 170, loss: 0.014730680733919144
step: 180, loss: 0.05496084690093994
step: 190, loss: 0.06980933994054794
step: 200, loss: 0.0003061093739233911
step: 210, loss: 0.0004922673106193542
step: 220, loss: 0.0010589660378172994
step: 230, loss: 0.0908600240945816
/usr/local/lib/python3.7/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=0.967452300785634,
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: 0.007723123766481876
step: 10, loss: 0.0005137314437888563
step: 20, loss: 0.0032699613366276026
step: 30, loss: 0.002399869728833437

```

```

step: 40, loss: 0.0012799170799553394
step: 50, loss: 0.001106968498788774
step: 60, loss: 0.004230877384543419
step: 70, loss: 0.006149178836494684
step: 80, loss: 0.0014120384585112333
step: 90, loss: 0.0015380405820906162
step: 100, loss: 0.000505371019244194
step: 110, loss: 0.00017354084411635995
step: 120, loss: 0.00021571613615378737
step: 130, loss: 0.0002089899790007621
step: 140, loss: 0.00018867829930968583
step: 150, loss: 0.00032768052187748253
step: 160, loss: 0.03891143947839737
step: 170, loss: 0.0006922144675627351
step: 180, loss: 0.0014044004492461681
step: 190, loss: 0.0006649699644185603
step: 200, loss: 0.004238834138959646
step: 210, loss: 0.000346022832673043
step: 220, loss: 0.0003877574345096946
step: 230, loss: 0.0003502405306790024
/usr/local/lib/python3.7/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=0.971815107102593, f1=0.9642857142857144,
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: 0.0002640104212332517
step: 10, loss: 0.0004599575186148286
step: 20, loss: 0.00015282363165169954
step: 30, loss: 0.00017096269584726542
step: 40, loss: 0.0016660018591210246
step: 50, loss: 0.0024245737586170435
step: 60, loss: 0.0004002189380116761
step: 70, loss: 0.000574981328099966
step: 80, loss: 0.0003363689756952226
step: 90, loss: 0.000370873516658321
step: 100, loss: 0.0004554196202661842
step: 110, loss: 0.030712099745869637
step: 120, loss: 0.005222718697041273
step: 130, loss: 0.00020582716388162225
step: 140, loss: 0.0003411401412449777
step: 150, loss: 0.0007278675329871476
step: 160, loss: 0.000693178444635123
step: 170, loss: 0.0020069656893610954
step: 180, loss: 0.002248372882604599
step: 190, loss: 0.005698119290173054

```

```

step: 200, loss: 0.001121897017583251
step: 210, loss: 0.013062468729913235
step: 220, loss: 0.008130480535328388
step: 230, loss: 0.00048407536814920604
/usr/local/lib/python3.7/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=0.9663677130044843, f1=0.9652855543113102,
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: 0.0004459308402147144
step: 10, loss: 0.0008910122560337186
step: 20, loss: 0.0004981684032827616
step: 30, loss: 0.00016456547018606216
step: 40, loss: 0.015984494239091873
step: 50, loss: 0.00016438739839941263
step: 60, loss: 0.0006324811838567257
step: 70, loss: 0.00024359545204788446
step: 80, loss: 0.00029621145222336054
step: 90, loss: 0.000471437320811674
step: 100, loss: 0.0001402506313752383
step: 110, loss: 7.136404019547626e-05
step: 120, loss: 6.081740866648033e-05
step: 130, loss: 7.100459333742037e-05
step: 140, loss: 0.00016343149763997644
step: 150, loss: 0.007013535127043724
step: 160, loss: 0.000367784989066422
step: 170, loss: 0.0005733614671044052
step: 180, loss: 0.0015353224007412791
step: 190, loss: 0.0011442756513133645
step: 200, loss: 0.004505376797169447
step: 210, loss: 0.00039843094418756664
step: 220, loss: 0.0003667129494715482
step: 230, loss: 0.0018496447009965777
/usr/local/lib/python3.7/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=0.9627118644067796, f1=0.9619686800894856,
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: 0.006077237892895937
step: 10, loss: 4.282246300135739e-05
step: 20, loss: 0.0003629799175541848
step: 30, loss: 0.00010069936251966283

```

```

step: 40, loss: 0.000624754058662802
step: 50, loss: 0.020585864782333374
step: 60, loss: 0.009427233599126339
step: 70, loss: 0.0002897621598094702
step: 80, loss: 0.00012812820204999298
step: 90, loss: 0.001003130222670734
step: 100, loss: 0.0003515472635626793
step: 110, loss: 0.00022536053438670933
step: 120, loss: 0.000308350776322186
step: 130, loss: 0.0001521651865914464
step: 140, loss: 0.14156359434127808
step: 150, loss: 0.001988019561395049
step: 160, loss: 0.00031556037720292807
step: 170, loss: 0.0002284057263750583
step: 180, loss: 0.0012674498138949275
step: 190, loss: 0.000868332339450717
step: 200, loss: 0.00011665540660033002
step: 210, loss: 0.0003537542943377048
step: 220, loss: 0.00015558292216155678
step: 230, loss: 0.0002788772981148213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9730941704035874, f1=0.9665178571428571,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017030004528351128
step: 10, loss: 0.00023842962400522083
step: 20, loss: 0.001713883480988443
step: 30, loss: 0.001153363031335175
step: 40, loss: 0.004348310641944408
step: 50, loss: 0.00039090286009013653
step: 60, loss: 0.0003587358514778316
step: 70, loss: 9.907613275572658e-05
step: 80, loss: 8.38916894281283e-05
step: 90, loss: 6.437191768782213e-05
step: 100, loss: 8.225362398661673e-05
step: 110, loss: 0.004257523454725742
step: 120, loss: 0.0002352349110879004
step: 130, loss: 8.512751082889736e-05
step: 140, loss: 0.00011704112694133073
step: 150, loss: 0.027655940502882004

```



```

step: 160, loss: 5.418476939667016e-05
step: 170, loss: 0.002356015145778656
step: 180, loss: 0.0001042915610014461
step: 190, loss: 4.896125756204128e-05
step: 200, loss: 0.00804123654961586
step: 210, loss: 7.584123522974551e-05
step: 220, loss: 7.60011316742748e-05
step: 230, loss: 0.00010624439164530486
/usr/local/lib/python3.7/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=0.9709821428571428, f1=0.9688888888888889,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012226813123561442
step: 10, loss: 5.196837446419522e-05
step: 20, loss: 6.549262616317719e-05
step: 30, loss: 0.07563123106956482
step: 40, loss: 0.10069742798805237
step: 50, loss: 0.0011240638559684157
step: 60, loss: 0.001253784168511629
step: 70, loss: 0.00037281360710039735
step: 80, loss: 0.0002492171188350767
step: 90, loss: 5.836330456077121e-05
step: 100, loss: 0.0008114710217341781
step: 110, loss: 0.00010043055954156443
step: 120, loss: 0.0001888629049062729
step: 130, loss: 7.799797458574176e-05
step: 140, loss: 0.01755840703845024
step: 150, loss: 8.74441975611262e-05
step: 160, loss: 0.00013035540177952498
step: 170, loss: 0.0002489285543560982
step: 180, loss: 0.01778586208820343
step: 190, loss: 0.0004961126251146197
step: 200, loss: 7.203467248473316e-05
step: 210, loss: 8.93826800165698e-05
step: 220, loss: 5.22269110661e-05
step: 230, loss: 0.012588047422468662
/usr/local/lib/python3.7/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=0.9709821428571428, f1=0.9732142857142857,
best_f1=0.9665178571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00014219032891560346
step: 10, loss: 0.00032578568789176643
step: 20, loss: 0.0004779742448590696
step: 30, loss: 0.00010954052413580939
step: 40, loss: 0.0009094499982893467
step: 50, loss: 0.0004981684032827616
step: 60, loss: 0.00014645190094597638
step: 70, loss: 5.5133761634351686e-05
step: 80, loss: 0.00012871874787379056
step: 90, loss: 0.00022089671983849257
step: 100, loss: 0.00011408278078306466
step: 110, loss: 0.02608230523765087
step: 120, loss: 0.00296822190284729
step: 130, loss: 6.994720024522394e-05
step: 140, loss: 0.0005194912082515657
step: 150, loss: 9.206045069731772e-05
step: 160, loss: 0.0001019654591800645
step: 170, loss: 0.00018997187726199627
step: 180, loss: 8.165314648067579e-05
step: 190, loss: 8.848056313581765e-05
step: 200, loss: 0.0001990756281884387
step: 210, loss: 3.954118801630102e-05
step: 220, loss: 4.660409103962593e-05
step: 230, loss: 0.0014095902442932129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9731543624161074, f1=0.972129319955407,
best_f1=0.972129319955407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021838067914359272
step: 10, loss: 7.048173574730754e-05
step: 20, loss: 8.21590147097595e-05
step: 30, loss: 7.418014138238505e-05
step: 40, loss: 0.0029818136245012283
step: 50, loss: 8.761577191762626e-05
step: 60, loss: 4.132583489990793e-05
step: 70, loss: 8.176614210242406e-05
step: 80, loss: 5.136126492288895e-05
step: 90, loss: 7.036420720396563e-05
step: 100, loss: 5.425063500297256e-05
step: 110, loss: 0.00013831813703291118

```

```

step: 120, loss: 4.1513878386467695e-05
step: 130, loss: 0.00023285004135686904
step: 140, loss: 6.589859549421817e-05
step: 150, loss: 6.614611629629508e-05
step: 160, loss: 0.00016559792857151479
step: 170, loss: 0.0008234684355556965
step: 180, loss: 0.00016502045036759228
step: 190, loss: 9.139356552623212e-05
step: 200, loss: 0.014088200405240059
step: 210, loss: 5.853069887962192e-05
step: 220, loss: 9.50125468079932e-05
step: 230, loss: 0.0007562385872006416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9742441209406495, f1=0.9710467706013363,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.19206046697218e-05
step: 10, loss: 3.7517100281547755e-05
step: 20, loss: 7.227284368127584e-05
step: 30, loss: 4.8984096792992204e-05
step: 40, loss: 0.000554995087441057
step: 50, loss: 0.012304098345339298
step: 60, loss: 5.9413592680357397e-05
step: 70, loss: 8.594730024924502e-05
step: 80, loss: 0.00013538786151912063
step: 90, loss: 0.0001331127277808264
step: 100, loss: 4.6436485718004405e-05
step: 110, loss: 5.096406675875187e-05
step: 120, loss: 5.783807864645496e-05
step: 130, loss: 0.002801834372803569
step: 140, loss: 0.000407418207032606
step: 150, loss: 0.00016813409456517547
step: 160, loss: 2.774481072265189e-05
step: 170, loss: 3.56487107637804e-05
step: 180, loss: 0.0002769895363599062
step: 190, loss: 0.00011237180297030136
step: 200, loss: 5.2927600336261094e-05
step: 210, loss: 0.00019126664847135544
step: 220, loss: 0.00014384149108082056
step: 230, loss: 0.00011808740964625031

```

```

/usr/local/lib/python3.7/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=0.9727891156462585, f1=0.9718785151856018,
best_f1=0.9710467706013363
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, 201.29it/s]
load_f1 = 0.9738933030646991
real_f1 = 0.971687429218573
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, 232.89it/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: 0.6302250027656555
step: 10, loss: 0.6038451790809631
step: 20, loss: 0.5161104798316956
step: 30, loss: 0.09748882055282593
step: 40, loss: 0.2806627154350281
step: 50, loss: 0.2453959733247757
step: 60, loss: 0.12035278975963593
step: 70, loss: 0.11354031413793564
step: 80, loss: 0.07295063883066177
step: 90, loss: 0.28525063395500183
step: 100, loss: 0.16525551676750183
step: 110, loss: 0.05724275857210159
step: 120, loss: 0.0814087763428688
step: 130, loss: 0.05201048031449318
step: 140, loss: 0.10093515366315842
step: 150, loss: 0.06478074938058853
step: 160, loss: 0.048340797424316406
step: 170, loss: 0.29047518968582153
step: 180, loss: 0.049302928149700165
step: 190, loss: 0.019503595307469368
step: 200, loss: 0.11769647896289825
step: 210, loss: 0.09026382118463516
step: 220, loss: 0.22211149334907532
step: 230, loss: 0.1651073694229126
step: 240, loss: 0.041513584554195404
step: 250, loss: 0.028243031352758408

```

```

step: 260, loss: 0.050888728350400925
step: 270, loss: 0.017366793006658554
step: 280, loss: 0.053321704268455505
step: 290, loss: 0.10532017797231674
step: 300, loss: 0.11266354471445084
step: 310, loss: 0.31819409132003784
step: 320, loss: 0.15257909893989563
step: 330, loss: 0.16040053963661194
step: 340, loss: 0.09763430804014206
step: 350, loss: 0.013851833529770374
step: 360, loss: 0.21912352740764618
step: 370, loss: 0.22137920558452606
step: 380, loss: 0.058818988502025604
step: 390, loss: 0.12519849836826324
step: 400, loss: 0.30948740243911743
step: 410, loss: 0.07181332260370255
step: 420, loss: 0.04059557989239693
step: 430, loss: 0.18530656397342682
step: 440, loss: 0.099912129342556
step: 450, loss: 0.025031598284840584
step: 460, loss: 0.008938023820519447
step: 470, loss: 0.18649184703826904
step: 480, loss: 0.07858013361692429
step: 490, loss: 0.03912341967225075
step: 500, loss: 0.15591290593147278
step: 510, loss: 0.0549447238445282
step: 520, loss: 0.07969561219215393
step: 530, loss: 0.008781654760241508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9187411930483794, f1=0.9195509822263798,
best_f1=0.9195509822263798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08876228332519531
step: 10, loss: 0.04319675639271736
step: 20, loss: 0.11850115656852722
step: 30, loss: 0.016308004036545753
step: 40, loss: 0.14535073935985565
step: 50, loss: 0.17103733122348785
step: 60, loss: 0.004638063721358776
step: 70, loss: 0.043776825070381165

```

step: 80, loss: 0.048565544188022614
step: 90, loss: 0.017754605039954185
step: 100, loss: 0.06382699310779572
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 110, loss: 0.10140970349311829
step: 120, loss: 0.05352018028497696
step: 130, loss: 0.08571358025074005
step: 140, loss: 0.13956420123577118
step: 150, loss: 0.03773575648665428
step: 160, loss: 0.04560444876551628
step: 170, loss: 0.13334132730960846
step: 180, loss: 0.101840078830719
step: 190, loss: 0.07128775119781494
step: 200, loss: 0.01245121844112873
step: 210, loss: 0.029691461473703384
step: 220, loss: 0.06520841270685196
step: 230, loss: 0.03965844586491585
step: 240, loss: 0.009655789472162724
step: 250, loss: 0.07626362890005112
step: 260, loss: 0.01933286525309086
step: 270, loss: 0.11091397702693939
step: 280, loss: 0.093777135014534
step: 290, loss: 0.032721828669309616
step: 300, loss: 0.16835632920265198
step: 310, loss: 0.008610650897026062
step: 320, loss: 0.15062308311462402
step: 330, loss: 0.03260566294193268
step: 340, loss: 0.04457272216677666
step: 350, loss: 0.003002872923389077
step: 360, loss: 0.06869805604219437
step: 370, loss: 0.15922288596630096
step: 380, loss: 0.0481741838157177
step: 390, loss: 0.05202513560652733
step: 400, loss: 0.06318418681621552
step: 410, loss: 0.026560330763459206
step: 420, loss: 0.0314779095351696
step: 430, loss: 0.07775139808654785
step: 440, loss: 0.1117905005812645
step: 450, loss: 0.08316381275653839
step: 460, loss: 0.0516340434551239
step: 470, loss: 0.06466436386108398
step: 480, loss: 0.1486876904964447
step: 490, loss: 0.04633941501379013
step: 500, loss: 0.1712951511144638
step: 510, loss: 0.04505280777812004
step: 520, loss: 0.07663647830486298
step: 530, loss: 0.043983567506074905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:


```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.924119241192412, f1=0.9213993639254885,
best_f1=0.9213993639254885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1760009527206421
step: 10, loss: 0.07405032217502594
step: 20, loss: 0.12596289813518524
step: 30, loss: 0.14265458285808563
step: 40, loss: 0.05463588237762451
step: 50, loss: 0.15679363906383514
step: 60, loss: 0.10502288490533829
step: 70, loss: 0.004521934315562248
step: 80, loss: 0.003046893747523427
step: 90, loss: 0.004370460286736488
step: 100, loss: 0.05257042497396469
step: 110, loss: 0.02574666403234005
step: 120, loss: 0.025660738348960876
step: 130, loss: 0.012615772895514965
step: 140, loss: 0.09793974459171295
step: 150, loss: 0.010956361889839172
step: 160, loss: 0.02142884023487568
step: 170, loss: 0.06635164469480515
step: 180, loss: 0.018128501251339912
step: 190, loss: 0.0413098968565464
step: 200, loss: 0.023650797083973885
step: 210, loss: 0.05601651594042778
step: 220, loss: 0.011896486394107342
step: 230, loss: 0.12688812613487244
step: 240, loss: 0.007060409989207983
step: 250, loss: 0.11991072446107864
step: 260, loss: 0.12654545903205872
step: 270, loss: 0.02026602439582348
step: 280, loss: 0.041101280599832535
step: 290, loss: 0.002813482191413641
step: 300, loss: 0.06399372965097427
step: 310, loss: 0.06683029979467392
step: 320, loss: 0.0649319738149643
step: 330, loss: 0.024575350806117058
step: 340, loss: 0.07797335088253021
step: 350, loss: 0.006113077979534864
step: 360, loss: 0.08218758553266525

```

```

step: 370, loss: 0.027509603649377823
step: 380, loss: 0.028196679428219795
step: 390, loss: 0.006157590541988611
step: 400, loss: 0.031093386933207512
step: 410, loss: 0.0024311791639775038
step: 420, loss: 0.19460844993591309
step: 430, loss: 0.11904307454824448
step: 440, loss: 0.0024620427284389734
step: 450, loss: 0.0751820057630539
step: 460, loss: 0.019387628883123398
step: 470, loss: 0.0710376650094986
step: 480, loss: 0.012079360894858837
step: 490, loss: 0.0054804477840662
step: 500, loss: 0.0052123055793344975
step: 510, loss: 0.08942129462957382
step: 520, loss: 0.08144115656614304
step: 530, loss: 0.07565817981958389
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9325323475046211, f1=0.9212344541685858,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12098387628793716
step: 10, loss: 0.011998728848993778
step: 20, loss: 0.001081200665794313
step: 30, loss: 0.0008126205066218972
step: 40, loss: 0.001451744232326746
step: 50, loss: 0.15247711539268494
step: 60, loss: 0.011494290083646774
step: 70, loss: 0.0021878888364881277
step: 80, loss: 0.002061365405097604
step: 90, loss: 0.011755391955375671
step: 100, loss: 0.0029999157413840294
step: 110, loss: 0.04007978364825249
step: 120, loss: 0.0006663057720288634
step: 130, loss: 0.006214340217411518
step: 140, loss: 0.013181421905755997
step: 150, loss: 0.0373711921274662
step: 160, loss: 0.03899335116147995
step: 170, loss: 0.07521740347146988
step: 180, loss: 0.0021598709281533957

```

```

step: 190, loss: 0.08878687769174576
step: 200, loss: 0.01233365572988987
step: 210, loss: 0.012305489741265774
step: 220, loss: 0.11788327991962433
step: 230, loss: 0.040291380137205124
step: 240, loss: 0.03261631727218628
step: 250, loss: 0.01837640441954136
step: 260, loss: 0.0022529815323650837
step: 270, loss: 0.01090586930513382
step: 280, loss: 0.24225464463233948
step: 290, loss: 0.018812544643878937
step: 300, loss: 0.0016524099046364427
step: 310, loss: 0.05470063537359238
step: 320, loss: 0.05278979241847992
step: 330, loss: 0.006443447899073362
step: 340, loss: 0.011309956200420856
step: 350, loss: 0.04902617633342743
step: 360, loss: 0.011352160945534706
step: 370, loss: 0.0020838938653469086
step: 380, loss: 0.03230440616607666
step: 390, loss: 0.02070094272494316
step: 400, loss: 0.21535159647464752
step: 410, loss: 0.0015099863521754742
step: 420, loss: 0.029712751507759094
step: 430, loss: 0.10224878787994385
step: 440, loss: 0.043208807706832886
step: 450, loss: 0.026257144287228584
step: 460, loss: 0.11492492258548737
step: 470, loss: 0.11560314893722534
step: 480, loss: 0.10541535168886185
step: 490, loss: 0.012905392795801163
step: 500, loss: 0.04742595925927162
step: 510, loss: 0.011798305436968803
step: 520, loss: 0.06797755509614944
step: 530, loss: 0.04671471565961838
/usr/local/lib/python3.7/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=0.9258384506376948, f1=0.9183577159037282,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.058638691902160645
step: 10, loss: 0.05339004099369049
step: 20, loss: 0.004188617691397667
step: 30, loss: 0.013172424398362637
step: 40, loss: 0.031248265877366066

```

step: 50, loss: 0.01692415215075016
step: 60, loss: 0.036754775792360306
step: 70, loss: 0.013767620548605919
step: 80, loss: 0.004325320012867451
step: 90, loss: 0.015936600044369698
step: 100, loss: 0.0008381104562431574
step: 110, loss: 0.0007968852878548205
step: 120, loss: 0.01982778124511242
step: 130, loss: 0.001191031071357429
step: 140, loss: 0.000769401784054935
step: 150, loss: 0.001801306032575667
step: 160, loss: 0.0006443828460760415
step: 170, loss: 0.03072969987988472
step: 180, loss: 0.006691726390272379
step: 190, loss: 0.00916751567274332
step: 200, loss: 0.00055363291176036
step: 210, loss: 0.05589164420962334
step: 220, loss: 0.01129551138728857
step: 230, loss: 0.04896091669797897
step: 240, loss: 0.027588438242673874
step: 250, loss: 0.004020960070192814
step: 260, loss: 0.029538594186306
step: 270, loss: 0.0007911713910289109
step: 280, loss: 0.005663261748850346
step: 290, loss: 0.15720519423484802
step: 300, loss: 0.02179023064672947
step: 310, loss: 0.015924178063869476
step: 320, loss: 0.06302438676357269
step: 330, loss: 0.0570971854031086
step: 340, loss: 0.005905468482524157
step: 350, loss: 0.0033506930340081453
step: 360, loss: 0.010528025217354298
step: 370, loss: 0.02127882093191147
step: 380, loss: 0.0008114184020087123
step: 390, loss: 0.002848926931619644
step: 400, loss: 0.0017209772486239672
step: 410, loss: 0.0036415723152458668
step: 420, loss: 0.00978479441255331
step: 430, loss: 0.004956972319632769
step: 440, loss: 0.006404751446098089
step: 450, loss: 0.011752728372812271
step: 460, loss: 0.0020956178195774555
step: 470, loss: 0.002014781581237912
step: 480, loss: 0.0110493004322052
step: 490, loss: 0.0007918477058410645
step: 500, loss: 0.012021144852042198
step: 510, loss: 0.030527126044034958
step: 520, loss: 0.014780278317630291

```

step: 530, loss: 0.007682028226554394
/usr/local/lib/python3.7/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=0.9212454212454212, f1=0.9151765245300322,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00874044094234705
step: 10, loss: 0.00201193755492568
step: 20, loss: 0.030092110857367516
step: 30, loss: 0.013223436661064625
step: 40, loss: 0.010905597358942032
step: 50, loss: 0.05359882116317749
step: 60, loss: 0.0014007508289068937
step: 70, loss: 0.022007299587130547
step: 80, loss: 0.03872787579894066
step: 90, loss: 0.004998335149139166
step: 100, loss: 0.004706377163529396
step: 110, loss: 0.015420286916196346
step: 120, loss: 0.02252179943025112
step: 130, loss: 0.006553218699991703
step: 140, loss: 0.0018456911202520132
step: 150, loss: 0.011072161607444286
step: 160, loss: 0.013158202171325684
step: 170, loss: 0.009014456532895565
step: 180, loss: 0.0017579549457877874
step: 190, loss: 0.05516933277249336
step: 200, loss: 0.0019895625300705433
step: 210, loss: 0.14412324130535126
step: 220, loss: 0.012215739116072655
step: 230, loss: 0.0018216597381979227
step: 240, loss: 0.0038908047135919333
step: 250, loss: 0.010050315409898758
step: 260, loss: 0.08862768858671188
step: 270, loss: 0.2288464605808258
step: 280, loss: 0.017138194292783737
step: 290, loss: 0.00032645996543578804
step: 300, loss: 0.0009880606085062027
step: 310, loss: 0.004474916961044073
step: 320, loss: 0.0010632158955559134
step: 330, loss: 0.0036238946486264467
step: 340, loss: 0.08090467751026154
step: 350, loss: 0.022007599472999573
step: 360, loss: 0.004710082430392504
step: 370, loss: 0.0060706231743097305
step: 380, loss: 0.0015068614156916738

```

```

step: 390, loss: 0.02429390512406826
step: 400, loss: 0.0003751316398847848
step: 410, loss: 0.0008251933613792062
step: 420, loss: 0.0011433014879003167
step: 430, loss: 0.0068066539242863655
step: 440, loss: 0.00020681307069025934
step: 450, loss: 0.0013968328712508082
step: 460, loss: 0.0005989199853502214
step: 470, loss: 0.0018502050079405308
step: 480, loss: 0.029339835047721863
step: 490, loss: 0.021835025399923325
step: 500, loss: 0.07965972274541855
step: 510, loss: 0.041801560670137405
step: 520, loss: 0.015844237059354782
step: 530, loss: 0.009707520715892315
/usr/local/lib/python3.7/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=0.9147072383586906, f1=0.9075011504832029,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002905701519921422
step: 10, loss: 0.04377870261669159
step: 20, loss: 0.0053194435313344
step: 30, loss: 0.008224772289395332
step: 40, loss: 0.0036933317314833403
step: 50, loss: 0.06073717772960663
step: 60, loss: 0.06994025409221649
step: 70, loss: 0.0021049880888313055
step: 80, loss: 0.11922919005155563
step: 90, loss: 0.00039139713044278324
step: 100, loss: 0.004184774588793516
step: 110, loss: 0.005365288816392422
step: 120, loss: 0.006214737426489592
step: 130, loss: 0.007298514246940613
step: 140, loss: 0.0003853141679428518
step: 150, loss: 0.0014340534107759595
step: 160, loss: 0.0004395207797642797
step: 170, loss: 0.013589242473244667
step: 180, loss: 0.0006418642587959766
step: 190, loss: 0.0017067530425265431
step: 200, loss: 0.0013009586837142706
step: 210, loss: 0.002488624071702361
step: 220, loss: 0.0008795188623480499
step: 230, loss: 0.0009242697269655764
step: 240, loss: 0.001791432499885559

```

```

step: 250, loss: 0.002051058690994978
step: 260, loss: 0.00389857217669487
step: 270, loss: 0.00029448344139382243
step: 280, loss: 0.000524416973348707
step: 290, loss: 0.0026340738404542208
step: 300, loss: 0.003574914764612913
step: 310, loss: 0.006511386949568987
step: 320, loss: 0.0018496590200811625
step: 330, loss: 0.006237942725419998
step: 340, loss: 0.1143079623579979
step: 350, loss: 0.0004261976864654571
step: 360, loss: 0.014561928808689117
step: 370, loss: 0.05151280388236046
step: 380, loss: 0.0027601453475654125
step: 390, loss: 0.015200838446617126
step: 400, loss: 0.013640262186527252
step: 410, loss: 0.0023828090634196997
step: 420, loss: 0.0021784405689686537
step: 430, loss: 0.0018883362645283341
step: 440, loss: 0.014714059419929981
step: 450, loss: 0.001249093096703291
step: 460, loss: 0.0007307646446861327
step: 470, loss: 0.001342179486528039
step: 480, loss: 0.14275559782981873
step: 490, loss: 0.002202223753556609
step: 500, loss: 0.0005050423787906766
step: 510, loss: 0.0007912747096270323
step: 520, loss: 0.13942719995975494
step: 530, loss: 0.0034127114340662956
/usr/local/lib/python3.7/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=0.9283402681460935, f1=0.9224376731301939,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04639667645096779
step: 10, loss: 0.04428839683532715
step: 20, loss: 0.0001700367429293692
step: 30, loss: 0.2563575506210327
step: 40, loss: 0.0003089072706643492
step: 50, loss: 0.01057679858058691
step: 60, loss: 0.0007076546316966414
step: 70, loss: 0.00036983322934247553
step: 80, loss: 0.0008833082392811775
step: 90, loss: 0.001209366717375815
step: 100, loss: 0.0007488750852644444

```

```
step: 110, loss: 0.0007314807153306901
step: 120, loss: 0.010682051070034504
step: 130, loss: 0.00034061583573929965
step: 140, loss: 0.0023006577976047993
step: 150, loss: 0.002977414056658745
step: 160, loss: 0.005348959006369114
step: 170, loss: 0.12049820274114609
step: 180, loss: 0.0003410865319892764
step: 190, loss: 0.0019568633288145065
step: 200, loss: 0.0010210213949903846
step: 210, loss: 0.0076292697340250015
step: 220, loss: 0.0025606707204133272
step: 230, loss: 0.007837476208806038
step: 240, loss: 0.005093211308121681
step: 250, loss: 0.0004048305272590369
step: 260, loss: 0.0014818183844909072
step: 270, loss: 0.008770718239247799
step: 280, loss: 0.030615665018558502
step: 290, loss: 0.0030398976523429155
step: 300, loss: 0.07608683407306671
step: 310, loss: 0.0029295324347913265
step: 320, loss: 0.0006802669959142804
step: 330, loss: 0.020052138715982437
step: 340, loss: 0.0026610333006829023
step: 350, loss: 0.0011559923877939582
step: 360, loss: 0.0013147371355444193
step: 370, loss: 0.0026985136792063713
step: 380, loss: 0.005854760762304068
step: 390, loss: 0.09775550663471222
step: 400, loss: 0.002515428001061082
step: 410, loss: 0.032712921500205994
step: 420, loss: 0.000891786243300885
step: 430, loss: 0.02569187618792057
step: 440, loss: 0.020303957164287567
step: 450, loss: 0.0027516174595803022
step: 460, loss: 0.0002328574628336355
step: 470, loss: 0.0030031688511371613
step: 480, loss: 0.06704086810350418
step: 490, loss: 0.016141269356012344
step: 500, loss: 0.00848457682877779
step: 510, loss: 0.09458467364311218
step: 520, loss: 0.0023805780801922083
step: 530, loss: 0.0028500198386609554
/usr/local/lib/python3.7/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=0.916400182731841, f1=0.9155963302752294,
best_f1=0.9212344541685858
```


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

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

step: 0, loss: 0.023266101256012917

step: 10, loss: 0.0002857996732927859

step: 20, loss: 0.012873944826424122

step: 30, loss: 0.0031101941131055355

step: 40, loss: 0.0082656005397439

step: 50, loss: 0.003032093634828925

step: 60, loss: 0.00041824011714197695

step: 70, loss: 0.0026342326309531927

step: 80, loss: 0.026598412543535233

step: 90, loss: 8.871910540619865e-05

step: 100, loss: 0.0006678003701381385

step: 110, loss: 0.00021668615227099508

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

step: 120, loss: 0.00029162713326513767

step: 130, loss: 0.010782865807414055

step: 140, loss: 0.0025916744489222765

step: 150, loss: 0.00046868156641721725

step: 160, loss: 0.00048773831804282963

step: 170, loss: 0.0019862456247210503

step: 180, loss: 0.0012362035922706127

step: 190, loss: 8.238603186327964e-05

step: 200, loss: 0.0014808750711381435

step: 210, loss: 0.0012722999090328813

step: 220, loss: 0.018710318952798843

step: 230, loss: 8.03864240879193e-05

step: 240, loss: 9.297644282924011e-05

step: 250, loss: 0.0012528634397312999

step: 260, loss: 0.04747233912348747

step: 270, loss: 7.40795730962418e-05

step: 280, loss: 6.229591963347048e-05

step: 290, loss: 0.0567450076341629

step: 300, loss: 0.0010603959672152996

step: 310, loss: 0.004587868694216013

step: 320, loss: 0.00013463989307638258

step: 330, loss: 0.0010963680688291788

step: 340, loss: 0.002952509792521596

step: 350, loss: 0.0017087379237636924

step: 360, loss: 8.62800152390264e-05

step: 370, loss: 0.014620008878409863

step: 380, loss: 0.0004372764378786087

step: 390, loss: 0.00012837632675655186

step: 400, loss: 0.1948128193616867

step: 410, loss: 0.00026787343085743487

step: 420, loss: 0.00023963910643942654

step: 430, loss: 7.071781146805733e-05

```

step: 440, loss: 0.14004188776016235
step: 450, loss: 0.000871015596203506
step: 460, loss: 0.04058503732085228
step: 470, loss: 0.0003485338238533586
step: 480, loss: 0.0002567890624050051
step: 490, loss: 0.006271893624216318
step: 500, loss: 0.0026048540603369474
step: 510, loss: 0.02865811623632908
step: 520, loss: 0.016452129930257797
step: 530, loss: 0.000902092142496258
/usr/local/lib/python3.7/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=0.9224952741020794, f1=0.9147727272727272,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04504238814115524
step: 10, loss: 0.00017232920799870044
step: 20, loss: 7.506286056013778e-05
step: 30, loss: 0.00015238046762533486
step: 40, loss: 0.0068795704282820225
step: 50, loss: 0.0007091272273100913
step: 60, loss: 0.00016458430036436766
step: 70, loss: 0.00020098606182727963
step: 80, loss: 0.0001166449292213656
step: 90, loss: 8.88704089447856e-05
step: 100, loss: 0.0003404493036214262
step: 110, loss: 0.0004080232174601406
step: 120, loss: 5.039574534748681e-05
step: 130, loss: 0.0003491508250590414
step: 140, loss: 0.00429672421887517
step: 150, loss: 8.05396048235707e-05
step: 160, loss: 0.00015649438137188554
step: 170, loss: 0.0006399747217074037
step: 180, loss: 0.0014831353910267353
step: 190, loss: 0.001790543901734054
step: 200, loss: 0.0020709228701889515
step: 210, loss: 0.00014434766490012407
step: 220, loss: 0.033473074436187744
step: 230, loss: 0.0062331752851605415
step: 240, loss: 0.02203412353992462
step: 250, loss: 0.00021854280203115195
step: 260, loss: 0.00019334979879204184
step: 270, loss: 0.005802854895591736
step: 280, loss: 5.132237492944114e-05
step: 290, loss: 0.0007258890545926988

```

```

step: 300, loss: 8.423512190347537e-05
step: 310, loss: 0.00036453717621043324
step: 320, loss: 0.002730043139308691
step: 330, loss: 0.007491314318031073
step: 340, loss: 8.661273022880778e-05
step: 350, loss: 0.17738023400306702
step: 360, loss: 0.0004222370916977525
step: 370, loss: 0.004112042486667633
step: 380, loss: 0.010753348469734192
step: 390, loss: 0.00013865919027011842
step: 400, loss: 0.0005203666514717042
step: 410, loss: 0.0022258548997342587
step: 420, loss: 0.0001913469604915008
step: 430, loss: 4.046557660331018e-05
step: 440, loss: 0.011149042285978794
step: 450, loss: 0.0004039178602397442
step: 460, loss: 0.00012075287668267265
step: 470, loss: 0.00037053506821393967
step: 480, loss: 0.0023434290196746588
step: 490, loss: 0.02406022697687149
step: 500, loss: 0.000727738777641207
step: 510, loss: 0.0022046815138310194
step: 520, loss: 0.002380875637754798
step: 530, loss: 0.0009850080823525786
/usr/local/lib/python3.7/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=0.929007992477668, f1=0.9220595181861123,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005943115102127194
step: 10, loss: 0.004176100715994835
step: 20, loss: 7.37801383365877e-05
step: 30, loss: 0.00018138060113415122
step: 40, loss: 0.005661665927618742
step: 50, loss: 0.021024765446782112
step: 60, loss: 0.012049095705151558
step: 70, loss: 0.00025990375434048474
step: 80, loss: 0.00015357561642304063
step: 90, loss: 0.00023726679501123726
step: 100, loss: 0.011510899290442467
step: 110, loss: 0.0005626757047139108
step: 120, loss: 0.0005477702361531556
step: 130, loss: 0.00016974573372863233
step: 140, loss: 0.009983189404010773
step: 150, loss: 4.1381957998964936e-05

```

```

step: 160, loss: 4.88298901473172e-05
step: 170, loss: 6.302051042439416e-05
step: 180, loss: 0.0019864484202116728
step: 190, loss: 0.00038629831396974623
step: 200, loss: 0.00813739188015461
step: 210, loss: 3.814632509602234e-05
step: 220, loss: 0.0007556685013696551
step: 230, loss: 8.56292899698019e-05
step: 240, loss: 0.0007263542502187192
step: 250, loss: 3.422522422624752e-05
step: 260, loss: 1.8045080651063472e-05
step: 270, loss: 8.533773507224396e-05
step: 280, loss: 0.0006629017880186439
step: 290, loss: 0.00010974716860800982
step: 300, loss: 0.0001675503299338743
step: 310, loss: 6.0189555370016024e-05
step: 320, loss: 0.00011244464985793456
step: 330, loss: 0.014465905725955963
step: 340, loss: 0.0011725407093763351
step: 350, loss: 0.08785989880561829
step: 360, loss: 0.00019493905710987747
step: 370, loss: 0.0024735284969210625
step: 380, loss: 0.002124931663274765
step: 390, loss: 0.001632816856727004
step: 400, loss: 0.0008110309136100113
step: 410, loss: 0.0002561317232903093
step: 420, loss: 6.419149576686323e-05
step: 430, loss: 0.00018253397138323635
step: 440, loss: 0.0002530553028918803
step: 450, loss: 0.00011380259093130007
step: 460, loss: 0.16269715130329132
step: 470, loss: 0.00018061599985230714
step: 480, loss: 0.0006868285126984119
step: 490, loss: 0.00019630380847956985
step: 500, loss: 0.00201388169080019
step: 510, loss: 0.00026346775121055543
step: 520, loss: 6.446083716582507e-05
step: 530, loss: 0.00025097018806263804
/usr/local/lib/python3.7/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=0.9202453987730062, f1=0.9183098591549296,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030986889032647014
step: 10, loss: 0.20935294032096863

```

step: 20, loss: 0.0016007269732654095
step: 30, loss: 0.0007365625351667404
step: 40, loss: 0.0015321398386731744
step: 50, loss: 0.0621129609644413
step: 60, loss: 0.0010023211361840367
step: 70, loss: 0.0006690490408800542
step: 80, loss: 0.0015578980091959238
step: 90, loss: 0.00013259038678370416
step: 100, loss: 0.0016761929728090763
step: 110, loss: 0.004970670212060213
step: 120, loss: 0.0005283395876176655
step: 130, loss: 0.00025706831365823746
step: 140, loss: 0.011514660902321339
step: 150, loss: 5.013977352064103e-05
step: 160, loss: 0.0008595040999352932
step: 170, loss: 0.0020603800658136606
step: 180, loss: 0.0011577350087463856
step: 190, loss: 0.001539932913146913
step: 200, loss: 6.995522562647238e-05
step: 210, loss: 0.0009122503688558936
step: 220, loss: 0.0004965708358213305
step: 230, loss: 3.1607782148057595e-05
step: 240, loss: 0.00070728495484218
step: 250, loss: 6.624701200053096e-05
step: 260, loss: 4.170846659690142e-05
step: 270, loss: 0.00013171069440431893
step: 280, loss: 0.0031224871054291725
step: 290, loss: 0.00010433148418087512
step: 300, loss: 0.0012554337736219168
step: 310, loss: 0.00012256018817424774
step: 320, loss: 0.00011402986274333671
step: 330, loss: 0.00011101758718723431
step: 340, loss: 5.6342956668231636e-05
step: 350, loss: 0.0002458595554344356
step: 360, loss: 0.013479214161634445
step: 370, loss: 0.00018090457888320088
step: 380, loss: 0.00022084172815084457
step: 390, loss: 0.005433692596852779
step: 400, loss: 0.00018017413094639778
step: 410, loss: 0.0006480916636064649
step: 420, loss: 0.007784045301377773
step: 430, loss: 0.0026068936567753553
step: 440, loss: 4.387778608361259e-05
step: 450, loss: 0.00039212327101267874
step: 460, loss: 0.011836127378046513
step: 470, loss: 2.3986329324543476e-05
step: 480, loss: 0.0001276315888389945
step: 490, loss: 0.020660489797592163

```

step: 500, loss: 0.00013708742335438728
step: 510, loss: 0.009684021584689617
step: 520, loss: 1.920340582728386e-05
step: 530, loss: 0.00565453851595521
/usr/local/lib/python3.7/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=0.9280806229958772, f1=0.923076923076923,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015003428561612964
step: 10, loss: 0.004843758884817362
step: 20, loss: 6.438112905016169e-05
step: 30, loss: 0.015284527093172073
step: 40, loss: 0.00023280714231077582
step: 50, loss: 0.0002563239831943065
step: 60, loss: 4.4482865632744506e-05
step: 70, loss: 0.0001271293731406331
step: 80, loss: 0.000892245676368475
step: 90, loss: 0.00010497234325157478
step: 100, loss: 0.025473689660429955
step: 110, loss: 0.00034878987935371697
step: 120, loss: 0.0003538321179803461
step: 130, loss: 8.154456736519933e-05
step: 140, loss: 0.00014352859579958022
step: 150, loss: 0.00019487336976453662
step: 160, loss: 3.0950384825700894e-05
step: 170, loss: 0.007285665720701218
step: 180, loss: 0.0002944829757325351
step: 190, loss: 0.0005169240175746381
step: 200, loss: 7.097709749359637e-05
step: 210, loss: 0.002658301964402199
step: 220, loss: 0.002350015565752983
step: 230, loss: 0.0018597589805722237
step: 240, loss: 0.0009411817300133407
step: 250, loss: 0.0002213567349826917
step: 260, loss: 0.00013640968245454133
step: 270, loss: 3.4993296139873564e-05
step: 280, loss: 0.00022547272965312004
step: 290, loss: 0.008919261395931244
step: 300, loss: 0.00018859909323509783
step: 310, loss: 0.001730430289171636
step: 320, loss: 0.0028710917104035616
step: 330, loss: 0.0017885328270494938
step: 340, loss: 0.0016336010303348303
step: 350, loss: 0.004935577046126127

```

```

step: 360, loss: 0.00024014578957576305
step: 370, loss: 0.0012134222779422998
step: 380, loss: 0.0001120098095270805
step: 390, loss: 0.000407145795179531
step: 400, loss: 0.000891119590960443
step: 410, loss: 0.0010353360557928681
step: 420, loss: 5.9812980907736346e-05
step: 430, loss: 0.0475681871175766
step: 440, loss: 0.004315834958106279
step: 450, loss: 0.000124356200103648
step: 460, loss: 0.0011796399485319853
step: 470, loss: 0.00014778849435970187
step: 480, loss: 0.002062553074210882
step: 490, loss: 0.005143217742443085
step: 500, loss: 0.00036132728564552963
step: 510, loss: 0.034076251089572906
step: 520, loss: 0.0001364851777907461
step: 530, loss: 0.00019268396135885268
/usr/local/lib/python3.7/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=0.928505957836847, f1=0.9248395967002749,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010587840806692839
step: 10, loss: 4.0428098145639524e-05
step: 20, loss: 4.213955980958417e-05
step: 30, loss: 0.00015189223631750792
step: 40, loss: 0.0004969554138369858
step: 50, loss: 0.0007784549961797893
step: 60, loss: 1.7083963030017912e-05
step: 70, loss: 0.00016067142132669687
step: 80, loss: 7.344842015299946e-05
step: 90, loss: 0.00040079187601804733
step: 100, loss: 0.0011265343055129051
step: 110, loss: 0.000255791557719931
step: 120, loss: 0.0005412030150182545
step: 130, loss: 0.029115255922079086
step: 140, loss: 0.0016571817686781287
step: 150, loss: 6.731542089255527e-05
step: 160, loss: 0.0005862706457264721
step: 170, loss: 0.00016616255743429065
step: 180, loss: 8.930249168770388e-05
step: 190, loss: 0.00018014713714364916
step: 200, loss: 0.00019093368609901518
step: 210, loss: 0.0061459727585315704

```

```

step: 220, loss: 0.000415733054978773
step: 230, loss: 0.006838504690676928
step: 240, loss: 7.648475730093196e-05
step: 250, loss: 6.137334275990725e-05
step: 260, loss: 0.0004901115316897631
step: 270, loss: 0.0002926003362517804
step: 280, loss: 4.104817708139308e-05
step: 290, loss: 0.08789726346731186
step: 300, loss: 0.0004964372492395341
step: 310, loss: 0.0086137093603611
step: 320, loss: 0.0005968902842141688
step: 330, loss: 8.889229502528906e-05
step: 340, loss: 0.0020492514595389366
step: 350, loss: 0.0004912580479867756
step: 360, loss: 0.00028543107328005135
step: 370, loss: 0.001385555020533502
step: 380, loss: 0.00426469836384058
step: 390, loss: 0.00392710929736495
step: 400, loss: 0.00014741692575626075
step: 410, loss: 5.7619814469944686e-05
step: 420, loss: 0.0004829696554224938
step: 430, loss: 0.00027551420498639345
step: 440, loss: 0.00012256513582542539
step: 450, loss: 0.0018081931630149484
step: 460, loss: 0.00014871392340864986
step: 470, loss: 0.0038288370706140995
step: 480, loss: 5.782215885119513e-05
step: 490, loss: 5.830790178151801e-05
step: 500, loss: 0.0005048480816185474
step: 510, loss: 0.0020215401891618967
step: 520, loss: 0.00030105069163255394
step: 530, loss: 0.00676416140049696
/usr/local/lib/python3.7/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=0.9251263206247129, f1=0.9236430542778289,
best_f1=0.9212344541685858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.23956070234999e-05
step: 10, loss: 0.0005285475053824484
step: 20, loss: 5.304102160152979e-05
step: 30, loss: 0.0002057532110484317
step: 40, loss: 0.024829963222146034
step: 50, loss: 0.013058536686003208
step: 60, loss: 0.00018493420793674886
step: 70, loss: 0.00038567109731957316

```


step: 80, loss: 7.870338595239446e-05
step: 90, loss: 7.101098162820563e-05
step: 100, loss: 0.005725916940718889
step: 110, loss: 0.0010993409669026732
step: 120, loss: 0.0004999577649869025
step: 130, loss: 0.06292437016963959
step: 140, loss: 0.00023989970213733613
step: 150, loss: 8.180052100215107e-05
step: 160, loss: 0.0003311385225970298
step: 170, loss: 0.047392960637807846
step: 180, loss: 0.0002764771634247154
step: 190, loss: 0.00010768631909741089
step: 200, loss: 0.0005052877822890878
step: 210, loss: 0.00022666405129712075
step: 220, loss: 0.0013387297512963414
step: 230, loss: 0.0007282859878614545
step: 240, loss: 0.0004138728545513004
step: 250, loss: 3.221457882318646e-05
step: 260, loss: 0.006094148848205805
step: 270, loss: 0.0008099455153569579
step: 280, loss: 0.00026352249551564455
step: 290, loss: 2.1870735508855432e-05
step: 300, loss: 4.12547196901869e-05
step: 310, loss: 0.0006411564536392689
step: 320, loss: 0.00027683531516231596
step: 330, loss: 0.00011553544027265161
step: 340, loss: 0.00020988352480344474
step: 350, loss: 4.2882740672212094e-05
step: 360, loss: 0.0024158854503184557
step: 370, loss: 4.505993638304062e-05
step: 380, loss: 0.0003833126393146813
step: 390, loss: 0.0010678964899852872
step: 400, loss: 0.0001541090023238212
step: 410, loss: 0.00023277146101463586
step: 420, loss: 0.0010536487679928541
step: 430, loss: 9.31398244574666e-05
step: 440, loss: 0.00040544537478126585
step: 450, loss: 4.3451200326671824e-05
step: 460, loss: 5.067072925157845e-05
step: 470, loss: 0.0036464217118918896
step: 480, loss: 1.836166666180361e-05
step: 490, loss: 1.7568238035892136e-05
step: 500, loss: 0.07054716348648071
step: 510, loss: 0.00011602212907746434
step: 520, loss: 6.377444515237585e-05
step: 530, loss: 7.842217746656388e-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=0.924860853432282, f1=0.9213691026827011,
best_f1=0.9212344541685858
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 232.73it/s]
load_f1 = 0.9300373134328358
real_f1 = 0.926852288815479
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, 247.98it/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: 0.587847888469696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.4, f1=0.3, best_f1=0.3
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4625902473926544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4545454545454545, f1=0.3703703703703704,
best_f1=0.3703703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42398425936698914
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4782608695652174, f1=0.44444444444444445,
best_f1=0.44444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24234354496002197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5454545454545454, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22795715928077698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.56, f1=0.5384615384615384, best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31284844875335693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.64, f1=0.5384615384615384, best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1613273024559021
/usr/local/lib/python3.7/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=0.5789473684210527, f1=0.5333333333333333,
best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0679832175374031
/usr/local/lib/python3.7/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=0.6153846153846153, f1=0.6153846153846153,
best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031599095091223717
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.689655172413793, f1=0.5714285714285714,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025994719937443733
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.6000000000000001,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00458888616412878
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.5806451612903226,
best_f1=0.5714285714285714

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003829779801890254
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.5384615384615384,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004520161543041468
/usr/local/lib/python3.7/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=0.6470588235294117, f1=0.5161290322580646,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003078499808907509
/usr/local/lib/python3.7/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=0.6470588235294117, f1=0.5161290322580646,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031352125108242035
/usr/local/lib/python3.7/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=0.6470588235294117, f1=0.5161290322580646,
best_f1=0.5714285714285714
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, 118424.35it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.5925925925925927
real_f1 = 0.6451612903225806
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.86it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

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

```



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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

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

step: 0, loss: 0.6242391467094421

```

step: 10, loss: 0.6266266107559204
step: 20, loss: 0.4266778826713562
step: 30, loss: 0.20987480878829956
step: 40, loss: 0.21948955953121185
step: 50, loss: 0.03903389722108841
step: 60, loss: 0.08414338529109955
step: 70, loss: 0.07976385951042175
step: 80, loss: 0.33827802538871765
step: 90, loss: 0.13062845170497894
step: 100, loss: 0.007839958183467388
step: 110, loss: 0.1834685504436493
step: 120, loss: 0.027192767709493637
step: 130, loss: 0.08174163848161697
step: 140, loss: 0.010036395862698555
step: 150, loss: 0.05165475234389305
step: 160, loss: 0.005022560711950064
step: 170, loss: 0.04914340004324913
step: 180, loss: 0.025765379890799522
step: 190, loss: 0.20060782134532928
step: 200, loss: 0.006420792546123266
step: 210, loss: 0.007259258534759283
step: 220, loss: 0.006005681585520506
step: 230, loss: 0.009157905355095863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9704545454545453, f1=0.9691428571428571,
best_f1=0.9691428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006734682247042656
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.005011027678847313
step: 20, loss: 0.18964609503746033
step: 30, loss: 0.20212820172309875
step: 40, loss: 0.024414876475930214
step: 50, loss: 0.003727738047018647
step: 60, loss: 0.029692169278860092
step: 70, loss: 0.15628230571746826
step: 80, loss: 0.0027548186480998993
step: 90, loss: 0.0179987121373415
step: 100, loss: 0.017782285809516907
step: 110, loss: 0.04863859340548515

```

```

step: 120, loss: 0.02533699944615364
step: 130, loss: 0.008118835277855396
step: 140, loss: 0.0026431081350892782
step: 150, loss: 0.0026938566006720066
step: 160, loss: 0.003721051849424839
step: 170, loss: 0.0019007865339517593
step: 180, loss: 0.00328909233212471
step: 190, loss: 0.025836337357759476
step: 200, loss: 0.018005317077040672
step: 210, loss: 0.0014629296492785215
step: 220, loss: 0.14592087268829346
step: 230, loss: 0.008364034816622734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798657718120806, f1=0.9731543624161074,
best_f1=0.9731543624161074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01884802244603634
step: 10, loss: 0.004649294540286064
step: 20, loss: 0.02532583847641945
step: 30, loss: 0.07122429460287094
step: 40, loss: 0.05856494605541229
step: 50, loss: 0.028740694746375084
step: 60, loss: 0.011538926512002945
step: 70, loss: 0.003411197802051902
step: 80, loss: 0.0015420462004840374
step: 90, loss: 0.13930128514766693
step: 100, loss: 0.0017618043348193169
step: 110, loss: 0.0065087322145700455
step: 120, loss: 0.010865382850170135
step: 130, loss: 0.018178531900048256
step: 140, loss: 0.01644333265721798
step: 150, loss: 0.0013578759972006083
step: 160, loss: 0.05789363756775856
step: 170, loss: 0.010008550249040127
step: 180, loss: 0.010929792188107967
step: 190, loss: 0.001306407735683024
step: 200, loss: 0.05442251265048981
step: 210, loss: 0.0016160471132025123
step: 220, loss: 0.0007812193944118917
step: 230, loss: 0.0007493476150557399

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854096520763187, f1=0.978675645342312,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005946848541498184
step: 10, loss: 0.0002395348419668153
step: 20, loss: 0.0026561312843114138
step: 30, loss: 0.00022864299535285681
step: 40, loss: 0.005014744587242603
step: 50, loss: 0.0017519722459837794
step: 60, loss: 0.000516207714099437
step: 70, loss: 0.00043679913505911827
step: 80, loss: 0.0007838273886591196
step: 90, loss: 0.0033117716666311026
step: 100, loss: 0.00987301953136921
step: 110, loss: 0.0003554584691300988
step: 120, loss: 0.069847971200943
step: 130, loss: 0.03356263041496277
step: 140, loss: 0.0039002487901598215
step: 150, loss: 0.05954599007964134
step: 160, loss: 0.0004829337412957102
step: 170, loss: 0.0038269897922873497
step: 180, loss: 0.000782207353040576
step: 190, loss: 0.005645770579576492
step: 200, loss: 0.0067988173104822636
step: 210, loss: 0.10684774070978165
step: 220, loss: 0.0001986689167097211
step: 230, loss: 0.10741663724184036
/usr/local/lib/python3.7/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=0.9772209567198178, f1=0.9761634506242906,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00043843622552230954
step: 10, loss: 0.0004450827545952052
step: 20, loss: 0.007556903641670942
step: 30, loss: 0.0031345109455287457

```

```

step: 40, loss: 0.0023928177542984486
step: 50, loss: 0.00033020967384800315
step: 60, loss: 0.00017091994232032448
step: 70, loss: 0.00012592741404660046
step: 80, loss: 0.00011333748261677101
step: 90, loss: 0.00011910550529137254
step: 100, loss: 0.00015466798504348844
step: 110, loss: 0.00015740703383926302
step: 120, loss: 0.00018963584443554282
step: 130, loss: 0.00037531161797232926
step: 140, loss: 0.002814473584294319
step: 150, loss: 0.0035831390414386988
step: 160, loss: 0.0003451882512308657
step: 170, loss: 0.014148014597594738
step: 180, loss: 0.0028120276983827353
step: 190, loss: 0.0005953506333753467
step: 200, loss: 0.0017511597834527493
step: 210, loss: 0.00019288867770228535
step: 220, loss: 0.0017949535977095366
step: 230, loss: 0.0007231712224893272
/usr/local/lib/python3.7/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=0.9819413092550789, f1=0.9775280898876404,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001273146364837885
step: 10, loss: 0.0011284974170848727
step: 20, loss: 0.002785699674859643
step: 30, loss: 0.002658666344359517
step: 40, loss: 0.00036400932003743947
step: 50, loss: 0.0006389129557646811
step: 60, loss: 0.0004940041108056903
step: 70, loss: 0.003793338779360056
step: 80, loss: 0.0004817110311705619
step: 90, loss: 0.0007529695867560804
step: 100, loss: 0.04906781390309334
step: 110, loss: 0.00028653652407228947
step: 120, loss: 0.0002506764139980078
step: 130, loss: 0.002290213480591774
step: 140, loss: 9.816742385737598e-05
step: 150, loss: 0.00034354435047134757
step: 160, loss: 0.0011052436893805861
step: 170, loss: 0.0020446819253265858
step: 180, loss: 0.007260940503329039
step: 190, loss: 0.07082823663949966

```

```

step: 200, loss: 0.0006163923535495996
step: 210, loss: 0.0004880149499513209
step: 220, loss: 0.00020276207942515612
step: 230, loss: 0.0023500409442931414
/usr/local/lib/python3.7/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=0.9822616407982262, f1=0.9722530521642618,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009309685789048672
step: 10, loss: 0.00018144186469726264
step: 20, loss: 0.00012862158473581076
step: 30, loss: 0.00016694888472557068
step: 40, loss: 0.00010027035023085773
step: 50, loss: 0.00011321876081638038
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.0005217221332713962
step: 70, loss: 0.0001745181652950123
step: 80, loss: 8.240487659350038e-05
step: 90, loss: 5.951894127065316e-05
step: 100, loss: 7.987471326487139e-05
step: 110, loss: 0.0029743502382189035
step: 120, loss: 0.0002692114794626832
step: 130, loss: 8.637017162982374e-05
step: 140, loss: 0.00011315302981529385
step: 150, loss: 0.0029333655256778
step: 160, loss: 0.19679898023605347
step: 170, loss: 0.021651841700077057
step: 180, loss: 0.0016980415675789118
step: 190, loss: 0.00012389480252750218
step: 200, loss: 0.03608980029821396
step: 210, loss: 0.00013511300494428724
step: 220, loss: 0.001460359781049192
step: 230, loss: 0.00027362830587662756
/usr/local/lib/python3.7/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=0.9832026875699889, f1=0.9752808988764046,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.54528441070579e-05
step: 10, loss: 0.00027005665469914675
step: 20, loss: 0.0009143012575805187

```

```

step: 30, loss: 9.01569874258712e-05
step: 40, loss: 0.0002694849972613156
step: 50, loss: 0.0030428965110331774
step: 60, loss: 5.887470251764171e-05
step: 70, loss: 0.00041880778735503554
step: 80, loss: 0.0001769327063811943
step: 90, loss: 5.2903156756656244e-05
step: 100, loss: 9.190425043925643e-05
step: 110, loss: 0.000942235579714179
step: 120, loss: 0.00011243653716519475
step: 130, loss: 7.205130532383919e-05
step: 140, loss: 5.5775060900487006e-05
step: 150, loss: 8.712062845006585e-05
step: 160, loss: 6.0613558162003756e-05
step: 170, loss: 7.841604383429512e-05
step: 180, loss: 9.572452836437151e-05
step: 190, loss: 0.0001878864277387038
step: 200, loss: 9.297556971432641e-05
step: 210, loss: 0.00015506109048146755
step: 220, loss: 8.219949813792482e-05
step: 230, loss: 7.545949483755976e-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 save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9854423292273236, f1=0.9753363228699552,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021428630861919373
step: 10, loss: 0.00010019232286140323
step: 20, loss: 0.0010927984258159995
step: 30, loss: 0.00040846530464477837
step: 40, loss: 0.02896260842680931
step: 50, loss: 0.00017212006787303835
step: 60, loss: 0.00010341432789573446
step: 70, loss: 0.0003498884616419673
step: 80, loss: 0.0002141281438525766
step: 90, loss: 0.00018773482588585466
step: 100, loss: 0.00010532359010539949
step: 110, loss: 6.473105167970061e-05
step: 120, loss: 0.00010345910413889214
step: 130, loss: 9.390605555381626e-05
step: 140, loss: 5.297815368976444e-05

```

```

step: 150, loss: 0.00025733542861416936
step: 160, loss: 5.916849841014482e-05
step: 170, loss: 0.0005734579754061997
step: 180, loss: 0.008062329143285751
step: 190, loss: 0.00041701726149767637
step: 200, loss: 0.0011711736442521214
step: 210, loss: 0.0006917942664586008
step: 220, loss: 5.600725853582844e-05
step: 230, loss: 0.00010797091817948967
/usr/local/lib/python3.7/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=0.9810479375696767, f1=0.9776785714285714,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005657189758494496
step: 10, loss: 6.631190626649186e-05
step: 20, loss: 5.879594755242579e-05
step: 30, loss: 0.0001246041210833937
step: 40, loss: 6.777016824344173e-05
step: 50, loss: 9.5322699053213e-05
step: 60, loss: 6.888202915433794e-05
step: 70, loss: 0.0006961021572351456
step: 80, loss: 6.12148578511551e-05
step: 90, loss: 0.00021262151130940765
step: 100, loss: 6.099780512158759e-05
step: 110, loss: 0.0006304444978013635
step: 120, loss: 0.006001182831823826
step: 130, loss: 0.014362671412527561
step: 140, loss: 0.00021584739442914724
step: 150, loss: 0.00766013702377677
step: 160, loss: 4.3665138946380466e-05
step: 170, loss: 5.2137780585326254e-05
step: 180, loss: 0.00021246541291475296
step: 190, loss: 0.005673495586961508
step: 200, loss: 6.458538700826466e-05
step: 210, loss: 0.00040924979839473963
step: 220, loss: 0.0003851281653624028
step: 230, loss: 0.14747408032417297
/usr/local/lib/python3.7/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=0.9807909604519773, f1=0.9763779527559054,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002835665945895016
step: 10, loss: 0.0003551999107003212
step: 20, loss: 0.0009087966755032539
step: 30, loss: 0.005054493900388479
step: 40, loss: 0.006894343998283148
step: 50, loss: 0.0005085880402475595
step: 60, loss: 0.0006411659414879978
step: 70, loss: 0.0007753423997201025
step: 80, loss: 6.157390453154221e-05
step: 90, loss: 0.0001418109895894304
step: 100, loss: 0.00010551218292675912
step: 110, loss: 6.306271825451404e-05
step: 120, loss: 6.356678204610944e-05
step: 130, loss: 0.0003809318004641682
step: 140, loss: 9.300765668740496e-05
step: 150, loss: 0.04026822745800018
step: 160, loss: 0.0002026966103585437
step: 170, loss: 0.02306022122502327
step: 180, loss: 0.00012595363659784198
step: 190, loss: 7.116395863704383e-05
step: 200, loss: 0.00013622864207718521
step: 210, loss: 6.798862887080759e-05
step: 220, loss: 5.192446769797243e-05
step: 230, loss: 7.989338337210938e-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=0.9820224719101124, f1=0.9776286353467561,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.106719906209037e-05
step: 10, loss: 0.0008192838868126273
step: 20, loss: 0.00020069845777470618
step: 30, loss: 8.02193972049281e-05
step: 40, loss: 0.0017333857249468565
step: 50, loss: 0.0006105411448515952
step: 60, loss: 0.0002027969603659585
step: 70, loss: 0.0009720682282932103
step: 80, loss: 9.761184628587216e-05
step: 90, loss: 0.0002806873235385865
step: 100, loss: 5.1907289162045345e-05
step: 110, loss: 0.00012767137377522886
step: 120, loss: 6.02511026954744e-05
step: 130, loss: 4.905588139081374e-05
step: 140, loss: 6.839012348791584e-05

```

```

step: 150, loss: 0.00015218163025565445
step: 160, loss: 7.563504914287478e-05
step: 170, loss: 7.498580089304596e-05
step: 180, loss: 0.0001313331740675494
step: 190, loss: 4.263297523721121e-05
step: 200, loss: 3.1354837119579315e-05
step: 210, loss: 7.171991455834359e-05
step: 220, loss: 6.256998312892392e-05
step: 230, loss: 0.09196766465902328
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.9774774774774775,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002935676311608404
step: 10, loss: 0.0002728900290094316
step: 20, loss: 0.00023785790835972875
step: 30, loss: 0.0001943271781783551
step: 40, loss: 0.00012857289402745664
step: 50, loss: 6.170063716126606e-05
step: 60, loss: 0.00011577363329706714
step: 70, loss: 3.620789721026085e-05
step: 80, loss: 4.668175461119972e-05
step: 90, loss: 0.0003917403519153595
step: 100, loss: 3.1093950383365154e-05
step: 110, loss: 0.020878398790955544
step: 120, loss: 0.031689129769802094
step: 130, loss: 5.998308188281953e-05
step: 140, loss: 4.0216782508650795e-05
step: 150, loss: 3.614160232245922e-05
step: 160, loss: 9.616209717933089e-05
step: 170, loss: 0.0002546050527598709
step: 180, loss: 0.00014580163406208158
step: 190, loss: 2.8829294024035335e-05
step: 200, loss: 0.00011674908455461264
step: 210, loss: 5.513658470590599e-05
step: 220, loss: 6.815020606154576e-05
step: 230, loss: 3.351144914631732e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9820224719101124, f1=0.978675645342312,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.625323006417602e-05
step: 10, loss: 0.0004734426620416343
step: 20, loss: 3.879715586663224e-05
step: 30, loss: 0.04048345237970352
step: 40, loss: 4.5732766011497006e-05
step: 50, loss: 0.0001706711482256651
step: 60, loss: 9.290183515986428e-05
step: 70, loss: 4.523789175436832e-05
step: 80, loss: 3.9751197618898004e-05
step: 90, loss: 3.8357644370989874e-05
step: 100, loss: 5.2922081522410735e-05
step: 110, loss: 4.0520899347029626e-05
step: 120, loss: 2.4914052119129337e-05
step: 130, loss: 3.806376116699539e-05
step: 140, loss: 8.173857349902391e-05
step: 150, loss: 4.991953755961731e-05
step: 160, loss: 0.0018807606538757682
step: 170, loss: 3.184967135894112e-05
step: 180, loss: 4.345334673416801e-05
step: 190, loss: 8.103580330498517e-05
step: 200, loss: 5.62880486540962e-05
step: 210, loss: 0.00014572619693353772
step: 220, loss: 4.594515485223383e-05
step: 230, loss: 9.359698742628098e-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=0.9799107142857142, f1=0.9733333333333333,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.2863474189070985e-05
step: 10, loss: 2.700745790207293e-05
step: 20, loss: 5.45789698662702e-05
step: 30, loss: 0.00017816766921896487
step: 40, loss: 6.657730409642681e-05
step: 50, loss: 0.0055821724236011505
step: 60, loss: 4.998300573788583e-05
step: 70, loss: 3.881930388160981e-05
step: 80, loss: 4.08367341151461e-05
step: 90, loss: 7.518009806517512e-05
step: 100, loss: 4.6973298594821244e-05
step: 110, loss: 4.151550092501566e-05
step: 120, loss: 0.00018712000746745616
step: 130, loss: 6.994367140578106e-05
step: 140, loss: 3.1265415600501e-05

```

```

step: 150, loss: 0.00021422161080408841
step: 160, loss: 3.1123570806812495e-05
step: 170, loss: 2.3807870093150996e-05
step: 180, loss: 0.005106108263134956
step: 190, loss: 0.00011858913057949394
step: 200, loss: 0.00013466682867147028
step: 210, loss: 3.1127561669563875e-05
step: 220, loss: 4.2443221900612116e-05
step: 230, loss: 4.13936359109357e-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=0.980963045912654, f1=0.9764837625979844,
best_f1=0.9753363228699552
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:15, 160.61it/s]
load_f1 = 0.9853768278965129
real_f1 = 0.9831649831649831
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.43it/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: 0.6250797510147095
step: 10, loss: 0.5887187123298645
step: 20, loss: 0.5641039609909058
step: 30, loss: 0.254226952791214
step: 40, loss: 0.18394486606121063
step: 50, loss: 0.16127072274684906
step: 60, loss: 0.05452246218919754
step: 70, loss: 0.11159274727106094
step: 80, loss: 0.06729274988174438
step: 90, loss: 0.1419660747051239
step: 100, loss: 0.11345984786748886
step: 110, loss: 0.08659932017326355
step: 120, loss: 0.06891027092933655
step: 130, loss: 0.03424646332859993
step: 140, loss: 0.18460920453071594
step: 150, loss: 0.08082921802997589
step: 160, loss: 0.05057058855891228
step: 170, loss: 0.23384569585323334
step: 180, loss: 0.17403142154216766
step: 190, loss: 0.00946894846856594
step: 200, loss: 0.19892019033432007
step: 210, loss: 0.04324018582701683
step: 220, loss: 0.14251936972141266

```

```

step: 230, loss: 0.1410512924194336
step: 240, loss: 0.20615412294864655
step: 250, loss: 0.042177148163318634
step: 260, loss: 0.07147485762834549
step: 270, loss: 0.02912270277738571
step: 280, loss: 0.15640710294246674
step: 290, loss: 0.048650048673152924
step: 300, loss: 0.05845653638243675
step: 310, loss: 0.3174642026424408
step: 320, loss: 0.11068795621395111
step: 330, loss: 0.09096035361289978
step: 340, loss: 0.06337586790323257
step: 350, loss: 0.21253876388072968
step: 360, loss: 0.044987402856349945
step: 370, loss: 0.060936544090509415
step: 380, loss: 0.04929341748356819
step: 390, loss: 0.14294859766960144
step: 400, loss: 0.280866801738739
step: 410, loss: 0.0502798818051815
step: 420, loss: 0.011007744818925858
step: 430, loss: 0.1935080885887146
step: 440, loss: 0.014311693608760834
step: 450, loss: 0.026838410645723343
step: 460, loss: 0.01501809898763895
step: 470, loss: 0.19534605741500854
step: 480, loss: 0.04733283072710037
step: 490, loss: 0.16334964334964752
step: 500, loss: 0.059830646961927414
step: 510, loss: 0.06238624453544617
step: 520, loss: 0.12384489923715591
step: 530, loss: 0.004485259298235178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.917883211678832, f1=0.912471655328798,
best_f1=0.912471655328798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05575508624315262
step: 10, loss: 0.04307973384857178
step: 20, loss: 0.015043281018733978
step: 30, loss: 0.1440950483083725
step: 40, loss: 0.0964432954788208

```

step: 50, loss: 0.23632508516311646
step: 60, loss: 0.01331079937517643
step: 70, loss: 0.02331741712987423
step: 80, loss: 0.03925775736570358
step: 90, loss: 0.012857876718044281
step: 100, loss: 0.09309569001197815
step: 110, loss: 0.038390807807445526
step: 120, loss: 0.08959243446588516
step: 130, loss: 0.16528062522411346
step: 140, loss: 0.1490229070186615
step: 150, loss: 0.06494206935167313
step: 160, loss: 0.03392515331506729
step: 170, loss: 0.06273775547742844
step: 180, loss: 0.04298676922917366
step: 190, loss: 0.09694219380617142
step: 200, loss: 0.07954411953687668
step: 210, loss: 0.0157123152166605
step: 220, loss: 0.06496986746788025
step: 230, loss: 0.014374416321516037
step: 240, loss: 0.12080831825733185
step: 250, loss: 0.01435163151472807
step: 260, loss: 0.003679788438603282
step: 270, loss: 0.30254876613616943
step: 280, loss: 0.030759580433368683
step: 290, loss: 0.02699473686516285
step: 300, loss: 0.09248959273099899
step: 310, loss: 0.01018295157700777
step: 320, loss: 0.10172227770090103
step: 330, loss: 0.027853596955537796
step: 340, loss: 0.007860050536692142
step: 350, loss: 0.0903649851679802
step: 360, loss: 0.23068548738956451
step: 370, loss: 0.14737568795681
step: 380, loss: 0.05501451715826988
step: 390, loss: 0.0870572179555893
step: 400, loss: 0.08602914214134216
step: 410, loss: 0.006495498586446047
step: 420, loss: 0.10560794919729233
step: 430, loss: 0.00508938729763031
step: 440, loss: 0.12025956064462662
step: 450, loss: 0.033345095813274384
step: 460, loss: 0.0577237568795681
step: 470, loss: 0.08972654491662979
step: 480, loss: 0.14948216080665588
step: 490, loss: 0.01971556432545185
step: 500, loss: 0.39314115047454834
step: 510, loss: 0.014877265319228172
step: 520, loss: 0.03733058273792267


```

step: 530, loss: 0.054137397557497025
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9264029438822446, f1=0.921451538814883,
best_f1=0.921451538814883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03350672125816345
step: 10, loss: 0.043358806520700455
step: 20, loss: 0.09821482747793198
step: 30, loss: 0.12902672588825226
step: 40, loss: 0.00540546141564846
step: 50, loss: 0.13534078001976013
step: 60, loss: 0.016144467517733574
step: 70, loss: 0.0058717066422104836
step: 80, loss: 0.0031798393465578556
step: 90, loss: 0.0042688134126365185
step: 100, loss: 0.043144553899765015
step: 110, loss: 0.018871426582336426
step: 120, loss: 0.00902608223259449
step: 130, loss: 0.007947873324155807
step: 140, loss: 0.030906593427062035
step: 150, loss: 0.11669730395078659
step: 160, loss: 0.11036603897809982
step: 170, loss: 0.12382383644580841
step: 180, loss: 0.07790783047676086
step: 190, loss: 0.061001796275377274
step: 200, loss: 0.018860001116991043
step: 210, loss: 0.042724233120679855
step: 220, loss: 0.03272182121872902
step: 230, loss: 0.1782936006784439
step: 240, loss: 0.016425805166363716
step: 250, loss: 0.050965528935194016
step: 260, loss: 0.010764403268694878
step: 270, loss: 0.005381994415074587
step: 280, loss: 0.16192898154258728
step: 290, loss: 0.015924474224448204
step: 300, loss: 0.04317039996385574
step: 310, loss: 0.022734813392162323
step: 320, loss: 0.022821839898824692
step: 330, loss: 0.002735887421295047
step: 340, loss: 0.004083527252078056

```

```

step: 350, loss: 0.011556323617696762
step: 360, loss: 0.039252039045095444
step: 370, loss: 0.003089030273258686
step: 380, loss: 0.00238172453828156
step: 390, loss: 0.003438380779698491
step: 400, loss: 0.017920251935720444
step: 410, loss: 0.012894213199615479
step: 420, loss: 0.1600913256406784
step: 430, loss: 0.0410740002989769
step: 440, loss: 0.0017713082488626242
step: 450, loss: 0.08658625930547714
step: 460, loss: 0.018259769305586815
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 470, loss: 0.10197696834802628
step: 480, loss: 0.002909349277615547
step: 490, loss: 0.00199108081869781
step: 500, loss: 0.006668124347925186
step: 510, loss: 0.0013921897625550628
step: 520, loss: 0.031242666766047478
step: 530, loss: 0.006094149313867092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9286367795059469, f1=0.9240681086056144,
best_f1=0.9240681086056144
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025600600987672806
step: 10, loss: 0.007161113899201155
step: 20, loss: 0.005896615330129862
step: 30, loss: 0.001799346529878676
step: 40, loss: 0.01949252188205719
step: 50, loss: 0.004759685136377811
step: 60, loss: 0.004541624337434769
step: 70, loss: 0.05139601603150368
step: 80, loss: 0.09737925976514816
step: 90, loss: 0.029389524832367897
step: 100, loss: 0.0029048144351691008
step: 110, loss: 0.02109166979789734
step: 120, loss: 0.07819315046072006
step: 130, loss: 0.0026675884146243334
step: 140, loss: 0.0014740137849003077
step: 150, loss: 0.001168881542980671

```

```

step: 160, loss: 0.05407945439219475
step: 170, loss: 0.0007988677243702114
step: 180, loss: 0.0007073352462612092
step: 190, loss: 0.005302552133798599
step: 200, loss: 0.009596772491931915
step: 210, loss: 0.10828985273838043
step: 220, loss: 0.02457197941839695
step: 230, loss: 0.029144002124667168
step: 240, loss: 0.011445720680058002
step: 250, loss: 0.0054041254334151745
step: 260, loss: 0.0011280931066721678
step: 270, loss: 0.0009744773269630969
step: 280, loss: 0.006387227214872837
step: 290, loss: 0.013363951817154884
step: 300, loss: 0.0013764805626124144
step: 310, loss: 0.01847042143344879
step: 320, loss: 0.047149863094091415
step: 330, loss: 0.0010966008994728327
step: 340, loss: 0.033693477511405945
step: 350, loss: 0.0025040870532393456
step: 360, loss: 0.013027782551944256
step: 370, loss: 0.0063424562104046345
step: 380, loss: 0.003128227312117815
step: 390, loss: 0.054766945540905
step: 400, loss: 0.002718186704441905
step: 410, loss: 0.004542618524283171
step: 420, loss: 0.026281608268618584
step: 430, loss: 0.09608111530542374
step: 440, loss: 0.010396978817880154
step: 450, loss: 0.008348627015948296
step: 460, loss: 0.03843431919813156
step: 470, loss: 0.008388269692659378
step: 480, loss: 0.08647622913122177
step: 490, loss: 0.04724844545125961
step: 500, loss: 0.0084998095408082
step: 510, loss: 0.054417531937360764
step: 520, loss: 0.10872946679592133
step: 530, loss: 0.027241822332143784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9301675977653632, f1=0.9206200093940817,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.04678723216056824
step: 10, loss: 0.0794137716293335
step: 20, loss: 0.002552502788603306
step: 30, loss: 0.015337117947638035
step: 40, loss: 0.0010740571888163686
step: 50, loss: 0.015589299611747265
step: 60, loss: 0.1786511391401291
step: 70, loss: 0.0031260009855031967
step: 80, loss: 0.002460130024701357
step: 90, loss: 0.0012757471995428205
step: 100, loss: 0.0011343442602083087
step: 110, loss: 0.0006459447322413325
step: 120, loss: 0.00028681635740213096
step: 130, loss: 0.0007757823332212865
step: 140, loss: 0.03126660734415054
step: 150, loss: 0.004438495263457298
step: 160, loss: 0.020635532215237617
step: 170, loss: 0.001968069700524211
step: 180, loss: 0.06545449793338776
step: 190, loss: 0.0007100109942257404
step: 200, loss: 0.00044499948853626847
step: 210, loss: 0.0008868767763487995
step: 220, loss: 0.0014533885987475514
step: 230, loss: 0.004593241959810257
step: 240, loss: 0.0010927114635705948
step: 250, loss: 0.004249231424182653
step: 260, loss: 0.0009696796769276261
step: 270, loss: 0.0018741838866844773
step: 280, loss: 0.05056661739945412
step: 290, loss: 0.0949772372841835
step: 300, loss: 0.012931006029248238
step: 310, loss: 0.0021907268092036247
step: 320, loss: 0.03754889592528343
step: 330, loss: 0.030360383912920952
step: 340, loss: 0.0014180365251377225
step: 350, loss: 0.014971042983233929
step: 360, loss: 0.002958579920232296
step: 370, loss: 0.007045377977192402
step: 380, loss: 0.0009864837629720569
step: 390, loss: 0.0008893618360161781
step: 400, loss: 0.006767706014215946
step: 410, loss: 0.006080777384340763
step: 420, loss: 0.0036097296979278326
step: 430, loss: 0.0009906936902552843
step: 440, loss: 0.10266191512346268
step: 450, loss: 0.08649110049009323
```

```

step: 460, loss: 0.000967235944699496
step: 470, loss: 0.004393104929476976
step: 480, loss: 0.0015203234506770968
step: 490, loss: 0.0010624362621456385
step: 500, loss: 0.05089491605758667
step: 510, loss: 0.019357478246092796
step: 520, loss: 0.030453043058514595
step: 530, loss: 0.0011756070889532566
/usr/local/lib/python3.7/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=0.928772258669166, f1=0.9231494578029231,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009808354079723358
step: 10, loss: 0.0004252186918165535
step: 20, loss: 0.14085093140602112
step: 30, loss: 0.001645568641833961
step: 40, loss: 0.00025147420819848776
step: 50, loss: 0.07115717977285385
step: 60, loss: 0.0004055105964653194
step: 70, loss: 0.0003944097552448511
step: 80, loss: 0.001841581892222166
step: 90, loss: 0.0021935799159109592
step: 100, loss: 0.000953779264818877
step: 110, loss: 0.002639350015670061
step: 120, loss: 0.009910915978252888
step: 130, loss: 0.0004932831507176161
step: 140, loss: 0.018125537782907486
step: 150, loss: 0.001288301544263959
step: 160, loss: 0.0020090951584279537
step: 170, loss: 0.003197686281055212
step: 180, loss: 0.0013022196944803
step: 190, loss: 0.0018449501367285848
step: 200, loss: 0.0012633450096473098
step: 210, loss: 0.0010828456142917275
step: 220, loss: 0.0006337978411465883
step: 230, loss: 0.13861051201820374
step: 240, loss: 0.08574697375297546
step: 250, loss: 0.0002463820856064558
step: 260, loss: 0.00016386697825510055
step: 270, loss: 0.052335191518068314
step: 280, loss: 0.0010386898647993803
step: 290, loss: 0.00034951764973811805
step: 300, loss: 0.0010704539017751813
step: 310, loss: 0.00046130508417263627

```

```

step: 320, loss: 0.004442573990672827
step: 330, loss: 0.0035402141511440277
step: 340, loss: 0.01443234272301197
step: 350, loss: 0.06911952048540115
step: 360, loss: 0.0028464454226195812
step: 370, loss: 0.006664742715656757
step: 380, loss: 0.0017402579542249441
step: 390, loss: 0.003874958259984851
step: 400, loss: 0.00039665820077061653
step: 410, loss: 0.00015473642270080745
step: 420, loss: 0.0003209371934644878
step: 430, loss: 0.00033329735742881894
step: 440, loss: 0.0006968402303755283
step: 450, loss: 0.0002613603719510138
step: 460, loss: 0.00027816632064059377
step: 470, loss: 0.05831604078412056
step: 480, loss: 0.006130289286375046
step: 490, loss: 0.004735032096505165
step: 500, loss: 0.00012573861749842763
step: 510, loss: 0.0009442645823583007
step: 520, loss: 0.0016253796638920903
step: 530, loss: 0.054195817559957504
/usr/local/lib/python3.7/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=0.921189240207645, f1=0.9158091674462113,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003558097407221794
step: 10, loss: 0.049276202917099
step: 20, loss: 0.0034421905875205994
step: 30, loss: 0.0007514334283769131
step: 40, loss: 0.0007709570345468819
step: 50, loss: 0.00173224660102278
step: 60, loss: 0.0005805365508422256
step: 70, loss: 0.014695494435727596
step: 80, loss: 0.0012292233295738697
step: 90, loss: 0.00025742221623659134
step: 100, loss: 9.151455742539838e-05
step: 110, loss: 8.801645890343934e-05
step: 120, loss: 5.617044371319935e-05
step: 130, loss: 0.0011019896483048797
step: 140, loss: 6.125242362031713e-05
step: 150, loss: 0.0015426294412463903
step: 160, loss: 0.002501194830983877
step: 170, loss: 0.00033446215093135834

```

```

step: 180, loss: 0.003721419954672456
step: 190, loss: 0.020262954756617546
step: 200, loss: 0.0012120881583541632
step: 210, loss: 0.0012567524099722505
step: 220, loss: 4.620674371835776e-05
step: 230, loss: 0.009867546148598194
step: 240, loss: 0.005841247737407684
step: 250, loss: 0.0023788856342434883
step: 260, loss: 4.5486689487006515e-05
step: 270, loss: 9.08671718207188e-05
step: 280, loss: 0.0001082979651982896
step: 290, loss: 6.887660856591538e-05
step: 300, loss: 0.00030479003908112645
step: 310, loss: 0.044493384659290314
step: 320, loss: 9.9831166153308e-05
step: 330, loss: 0.021962374448776245
step: 340, loss: 0.03671609237790108
step: 350, loss: 0.0037886332720518112
step: 360, loss: 0.0030612838454544544
step: 370, loss: 0.0006196188041940331
step: 380, loss: 0.009239790961146355
step: 390, loss: 0.0003866578917950392
step: 400, loss: 0.0053094299510121346
step: 410, loss: 0.00842411257326603
step: 420, loss: 0.008312501944601536
step: 430, loss: 0.00021985068451613188
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.00018194655422121286
step: 450, loss: 0.00047102911048568785
step: 460, loss: 0.00022373758838512003
step: 470, loss: 0.00590928690508008
step: 480, loss: 0.06678459793329239
step: 490, loss: 0.002481652656570077
step: 500, loss: 0.00034243465051986277
step: 510, loss: 0.0008732402347959578
step: 520, loss: 0.061503857374191284
step: 530, loss: 0.005253934301435947
/usr/local/lib/python3.7/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=0.9264432029795159, f1=0.9246511627906977,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010681685525923967
step: 10, loss: 0.019327916204929352
step: 20, loss: 0.00024009613844100386

```

step: 30, loss: 0.10002224147319794
step: 40, loss: 0.0005211925017647445
step: 50, loss: 0.0045039597898721695
step: 60, loss: 0.0017194128595292568
step: 70, loss: 7.954756438266486e-05
step: 80, loss: 0.003419416956603527
step: 90, loss: 0.0002956589451059699
step: 100, loss: 0.00470691779628396
step: 110, loss: 0.0007129745790734887
step: 120, loss: 0.0001772093091858551
step: 130, loss: 0.003558488329872489
step: 140, loss: 0.0011929775355383754
step: 150, loss: 0.002941492246463895
step: 160, loss: 0.0007233510841615498
step: 170, loss: 0.006228739861398935
step: 180, loss: 0.00014920582179911435
step: 190, loss: 0.004560065921396017
step: 200, loss: 0.001870226114988327
step: 210, loss: 0.005379363428801298
step: 220, loss: 0.003542690770700574
step: 230, loss: 0.008711464703083038
step: 240, loss: 0.0001434280857210979
step: 250, loss: 0.00018898029520642012
step: 260, loss: 6.898371793795377e-05
step: 270, loss: 0.0004364904307294637
step: 280, loss: 0.0017610914073884487
step: 290, loss: 0.0001185991641250439
step: 300, loss: 0.0004627776506822556
step: 310, loss: 0.00014240187010727823
step: 320, loss: 0.07187262922525406
step: 330, loss: 0.00014403574459720403
step: 340, loss: 0.00016933638835325837
step: 350, loss: 0.00048175553092733026
step: 360, loss: 0.0006620739004574716
step: 370, loss: 0.0008134344825521111
step: 380, loss: 0.0064505282789468765
step: 390, loss: 0.11817541718482971
step: 400, loss: 8.181968587450683e-05
step: 410, loss: 0.0011810759315267205
step: 420, loss: 0.0031540344934910536
step: 430, loss: 0.018356677144765854
step: 440, loss: 0.049425724893808365
step: 450, loss: 0.00023413210874423385
step: 460, loss: 0.00041438505286350846
step: 470, loss: 0.0006701119709759951
step: 480, loss: 0.0002686635416466743
step: 490, loss: 0.0005494492361322045
step: 500, loss: 0.00384723162278533


```

step: 510, loss: 0.0002472735068295151
step: 520, loss: 0.001934673055075109
step: 530, loss: 0.00011866749991895631
/usr/local/lib/python3.7/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=0.925891181988743, f1=0.9197559831065226,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018322638934478164
step: 10, loss: 7.32460684957914e-05
step: 20, loss: 0.0025750305503606796
step: 30, loss: 0.00017799035413190722
step: 40, loss: 5.8040219300892204e-05
step: 50, loss: 0.0008959707338362932
step: 60, loss: 0.00011272783012827858
step: 70, loss: 5.343021985027008e-05
step: 80, loss: 0.0012175935553386807
step: 90, loss: 4.817948865820654e-05
step: 100, loss: 0.0006645424873568118
step: 110, loss: 0.00015861117572057992
step: 120, loss: 0.0004600349930115044
step: 130, loss: 0.008751949295401573
step: 140, loss: 0.00033132240059785545
step: 150, loss: 6.767792365280911e-05
step: 160, loss: 0.0012173072900623083
step: 170, loss: 0.005235464312136173
step: 180, loss: 5.388877616496757e-05
step: 190, loss: 0.00012580177281051874
step: 200, loss: 0.0006382153951562941
step: 210, loss: 3.897122587659396e-05
step: 220, loss: 0.00017750260303728282
step: 230, loss: 3.933612970286049e-05
step: 240, loss: 7.523285603383556e-05
step: 250, loss: 0.003910810220986605
step: 260, loss: 0.00034600007347762585
step: 270, loss: 2.3911954485811293e-05
step: 280, loss: 6.801607378292829e-05
step: 290, loss: 0.15972809493541718
step: 300, loss: 3.6874487705063075e-05
step: 310, loss: 0.023187890648841858
step: 320, loss: 0.00018402919522486627
step: 330, loss: 0.0005695761647075415
step: 340, loss: 0.00012776489893440157
step: 350, loss: 0.00019346906628925353
step: 360, loss: 0.00014842880773358047

```

```

step: 370, loss: 0.0018816690426319838
step: 380, loss: 0.0027462656144052744
step: 390, loss: 3.483272303128615e-05
step: 400, loss: 0.0012335613137111068
step: 410, loss: 8.78600767464377e-05
step: 420, loss: 3.761177868000232e-05
step: 430, loss: 6.789965846110135e-05
step: 440, loss: 0.0011669030645862222
step: 450, loss: 5.064687866251916e-05
step: 460, loss: 5.04674062540289e-05
step: 470, loss: 0.00027950244839303195
step: 480, loss: 2.7487656552693807e-05
step: 490, loss: 0.0031056164298206568
step: 500, loss: 0.0021301929373294115
step: 510, loss: 4.602320768753998e-05
step: 520, loss: 7.827287481632084e-05
step: 530, loss: 0.000196693858015351
/usr/local/lib/python3.7/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=0.9287696577243294, f1=0.9280742459396751,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.280570515722502e-05
step: 10, loss: 1.941188020282425e-05
step: 20, loss: 2.053293792414479e-05
step: 30, loss: 2.3386837710859254e-05
step: 40, loss: 0.00026092250482179224
step: 50, loss: 0.05659900978207588
step: 60, loss: 8.020392124308273e-05
step: 70, loss: 1.731118572934065e-05
step: 80, loss: 2.655308344401419e-05
step: 90, loss: 1.8801256373990327e-05
step: 100, loss: 1.7869968360173516e-05
step: 110, loss: 1.4331028069136664e-05
step: 120, loss: 0.006422759499400854
step: 130, loss: 0.00011536549573065713
step: 140, loss: 0.006148782558739185
step: 150, loss: 4.032504148199223e-05
step: 160, loss: 0.0004001730994787067
step: 170, loss: 0.00015595208969898522
step: 180, loss: 3.946812284993939e-05
step: 190, loss: 0.0005406655254773796
step: 200, loss: 9.999216854339465e-05
step: 210, loss: 5.3858668252360076e-05
step: 220, loss: 6.096833021729253e-05

```

```

step: 230, loss: 0.0018148522358387709
step: 240, loss: 3.0818100640317425e-05
step: 250, loss: 2.0872372260782868e-05
step: 260, loss: 0.00032127872691489756
step: 270, loss: 3.2249168725684285e-05
step: 280, loss: 6.41747610643506e-05
step: 290, loss: 0.001617533271200955
step: 300, loss: 2.498835419828538e-05
step: 310, loss: 0.00013737908739130944
step: 320, loss: 0.00580306863412261
step: 330, loss: 2.5706973246997222e-05
step: 340, loss: 0.00012392154894769192
step: 350, loss: 6.615990423597395e-05
step: 360, loss: 0.00025394820841029286
step: 370, loss: 0.005366286262869835
step: 380, loss: 0.0006430480862036347
step: 390, loss: 8.684357453603297e-05
step: 400, loss: 0.00023228874488268048
step: 410, loss: 0.0020217266865074635
step: 420, loss: 2.669048262760043e-05
step: 430, loss: 2.091695751005318e-05
step: 440, loss: 0.002319267252460122
step: 450, loss: 5.0208604079671204e-05
step: 460, loss: 0.00015813414938747883
step: 470, loss: 6.083244807086885e-05
step: 480, loss: 2.4154247512342408e-05
step: 490, loss: 3.269115404691547e-05
step: 500, loss: 0.0023034897167235613
step: 510, loss: 1.8644768715603277e-05
step: 520, loss: 2.198233596573118e-05
step: 530, loss: 0.0010348965879529715
/usr/local/lib/python3.7/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=0.9261682242990654, f1=0.9260808926080892,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.464337351033464e-05
step: 10, loss: 0.00035198492696508765
step: 20, loss: 8.546280150767416e-05
step: 30, loss: 5.5291275202762336e-05
step: 40, loss: 0.0042645628564059734
step: 50, loss: 0.003049843944609165
step: 60, loss: 0.04738639295101166
step: 70, loss: 0.00019773990788962692
step: 80, loss: 7.309587090276182e-05

```

```
step: 90, loss: 0.04685129597783089
step: 100, loss: 0.0011344114318490028
step: 110, loss: 7.22767726983875e-05
step: 120, loss: 5.2957719162805006e-05
step: 130, loss: 0.00010243301949230954
step: 140, loss: 0.15743772685527802
step: 150, loss: 0.009928376413881779
step: 160, loss: 0.0045269387774169445
step: 170, loss: 0.00013291335199028254
step: 180, loss: 0.00016614531341474503
step: 190, loss: 0.0013954900205135345
step: 200, loss: 0.005682979244738817
step: 210, loss: 0.0015501350862905383
step: 220, loss: 0.001319295377470553
step: 230, loss: 7.520099461544305e-05
step: 240, loss: 0.00021431346249300987
step: 250, loss: 4.3935648136539385e-05
step: 260, loss: 4.322181484894827e-05
step: 270, loss: 8.111312490655109e-05
step: 280, loss: 0.0002545974566601217
step: 290, loss: 0.00011961490963585675
step: 300, loss: 0.00014126359019428492
step: 310, loss: 0.00024061588919721544
step: 320, loss: 0.00018603485659696162
step: 330, loss: 5.540153870242648e-05
step: 340, loss: 0.00026183240697719157
step: 350, loss: 0.00011255427671130747
step: 360, loss: 0.0006147646927274764
step: 370, loss: 0.0009131599217653275
step: 380, loss: 5.465500726131722e-05
step: 390, loss: 2.666813088580966e-05
step: 400, loss: 0.00011891303438460454
step: 410, loss: 6.741393735865131e-05
step: 420, loss: 0.0004338362778071314
step: 430, loss: 2.9946180802653544e-05
step: 440, loss: 0.05042971670627594
step: 450, loss: 7.29370949557051e-05
step: 460, loss: 0.000399062322685495
step: 470, loss: 0.00020445515110623091
step: 480, loss: 0.00010211890185019001
step: 490, loss: 0.001080626156181097
step: 500, loss: 0.0022543095983564854
step: 510, loss: 0.000629057758487761
step: 520, loss: 7.558259676443413e-05
step: 530, loss: 0.00012601594789884984
/usr/local/lib/python3.7/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=0.9238095238095239, f1=0.9135802469135803,  
best_f1=0.9206200093940817  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 2.9328239179449156e-05  
step: 10, loss: 2.1803567506140098e-05  
step: 20, loss: 9.131793194683269e-05  
step: 30, loss: 0.008962715975940228  
step: 40, loss: 2.347249392187223e-05  
step: 50, loss: 0.04392389953136444  
step: 60, loss: 0.003539496101438999  
step: 70, loss: 7.163918780861422e-05  
step: 80, loss: 0.00018449542403686792  
step: 90, loss: 3.6885070585412905e-05  
step: 100, loss: 0.00017176076653413475  
step: 110, loss: 0.02156962640583515  
step: 120, loss: 3.4601751394802704e-05  
step: 130, loss: 0.0009424585732631385  
step: 140, loss: 3.9404221752192825e-05  
step: 150, loss: 0.00012809262261725962  
step: 160, loss: 6.33262752671726e-05  
step: 170, loss: 0.0003017761337105185  
step: 180, loss: 0.0006991034606471658  
step: 190, loss: 0.0009499956504441798  
step: 200, loss: 3.75003473891411e-05  
step: 210, loss: 2.756245703494642e-05  
step: 220, loss: 0.00013074545131530613  
step: 230, loss: 2.221663271484431e-05  
step: 240, loss: 0.00012975769641343504  
step: 250, loss: 1.9680192053783685e-05  
step: 260, loss: 4.958543286193162e-05  
step: 270, loss: 0.22519567608833313  
step: 280, loss: 0.004756158217787743  
step: 290, loss: 0.0007305640610866249  
step: 300, loss: 0.0022171675227582455  
step: 310, loss: 0.0009207238326780498  
step: 320, loss: 0.0014140127459540963  
step: 330, loss: 0.00016537723422516137  
step: 340, loss: 4.013722718809731e-05  
step: 350, loss: 0.00012881914153695107  
step: 360, loss: 0.0014276469592005014  
step: 370, loss: 0.025622021406888962  
step: 380, loss: 4.973714385414496e-05  
step: 390, loss: 7.260053098434582e-05  
step: 400, loss: 2.4515615223208442e-05  
step: 410, loss: 0.0004349726950749755  
step: 420, loss: 0.0103434594348073
```

```

step: 430, loss: 0.008696286007761955
step: 440, loss: 0.0002660305181052536
step: 450, loss: 0.0006597579922527075
step: 460, loss: 0.000266455375822261
step: 470, loss: 2.916334051406011e-05
step: 480, loss: 0.0010487495455890894
step: 490, loss: 0.0002804962277878076
step: 500, loss: 2.7438742108643055e-05
step: 510, loss: 0.00010361581371398643
step: 520, loss: 7.54788561607711e-05
step: 530, loss: 3.387190372450277e-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=0.9264229523368811, f1=0.9216589861751152,
best_f1=0.9206200093940817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030044084414839745
step: 10, loss: 3.0176268410286866e-05
step: 20, loss: 3.38489880959969e-05
step: 30, loss: 3.394703526282683e-05
step: 40, loss: 2.826891068252735e-05
step: 50, loss: 0.002880196087062359
step: 60, loss: 4.2937128455378115e-05
step: 70, loss: 6.149071123218164e-05
step: 80, loss: 1.6767255146987736e-05
step: 90, loss: 0.010203626938164234
step: 100, loss: 0.003064661519601941
step: 110, loss: 2.083089748339262e-05
step: 120, loss: 1.9270499251433648e-05
step: 130, loss: 3.123718124697916e-05
step: 140, loss: 5.286977466312237e-05
step: 150, loss: 3.2971827749861404e-05
step: 160, loss: 2.2935935703571886e-05
step: 170, loss: 0.0002354475436732173
step: 180, loss: 3.545021536410786e-05
step: 190, loss: 0.0027327649295330048
step: 200, loss: 2.0049123122589663e-05
step: 210, loss: 1.9672745111165568e-05
step: 220, loss: 1.787349538062699e-05
step: 230, loss: 6.357634265441447e-05
step: 240, loss: 8.787531987763941e-05
step: 250, loss: 4.5763852540403605e-05
step: 260, loss: 1.4733316675119568e-05
step: 270, loss: 1.3630671674036421e-05
step: 280, loss: 2.201958341174759e-05

```

```

step: 290, loss: 2.285383379785344e-05
step: 300, loss: 2.4109360310831107e-05
step: 310, loss: 0.002224232302978635
step: 320, loss: 0.008574427105486393
step: 330, loss: 0.0018666093237698078
step: 340, loss: 5.091579077998176e-05
step: 350, loss: 7.990965968929231e-05
step: 360, loss: 2.05443830054719e-05
step: 370, loss: 3.230053698644042e-05
step: 380, loss: 2.367320666962769e-05
step: 390, loss: 7.648969040019438e-05
step: 400, loss: 1.55342131620273e-05
step: 410, loss: 0.0002971876529045403
step: 420, loss: 1.8220011043013074e-05
step: 430, loss: 0.0028848163783550262
step: 440, loss: 2.134878195647616e-05
step: 450, loss: 2.5174589609378017e-05
step: 460, loss: 0.00018817614181898534
step: 470, loss: 1.522494221717352e-05
step: 480, loss: 2.1206851670285687e-05
step: 490, loss: 1.779914236976765e-05
step: 500, loss: 1.1581802937143948e-05
step: 510, loss: 7.636936061317101e-05
step: 520, loss: 0.0005365345859900117
step: 530, loss: 1.7229191143997014e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9308755760368663, f1=0.9273897058823529,
best_f1=0.9273897058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005738034378737211
step: 10, loss: 1.2870736100012437e-05
step: 20, loss: 1.424534639227204e-05
step: 30, loss: 1.8707905837800354e-05
step: 40, loss: 2.6720606911112554e-05
step: 50, loss: 6.854740786366165e-05
step: 60, loss: 1.9862471162923612e-05
step: 70, loss: 0.00017036657663993537
step: 80, loss: 0.0001857751194620505
step: 90, loss: 2.365498585277237e-05
step: 100, loss: 1.7724580175126903e-05

```

```
step: 110, loss: 1.715463622531388e-05
step: 120, loss: 2.0790286725969054e-05
step: 130, loss: 0.030519628897309303
step: 140, loss: 1.3440698239719495e-05
step: 150, loss: 1.8733899196377024e-05
step: 160, loss: 1.9445620637270622e-05
step: 170, loss: 0.0010717620607465506
step: 180, loss: 4.832641570828855e-05
step: 190, loss: 2.1110181478434242e-05
step: 200, loss: 3.361606286489405e-05
step: 210, loss: 1.7404276150045916e-05
step: 220, loss: 2.0715684513561428e-05
step: 230, loss: 1.7028018191922456e-05
step: 240, loss: 3.6543297028401867e-05
step: 250, loss: 1.4200647456164006e-05
step: 260, loss: 1.4904713680152781e-05
step: 270, loss: 1.6722453437978402e-05
step: 280, loss: 1.7784192095859908e-05
step: 290, loss: 0.02804575115442276
step: 300, loss: 2.180349110858515e-05
step: 310, loss: 0.12597854435443878
step: 320, loss: 2.7446900276117958e-05
step: 330, loss: 2.4984634364955127e-05
step: 340, loss: 1.939727553690318e-05
step: 350, loss: 6.21384460828267e-05
step: 360, loss: 3.2784755603643134e-05
step: 370, loss: 1.9233362763770856e-05
step: 380, loss: 8.082168642431498e-05
step: 390, loss: 0.008119452744722366
step: 400, loss: 0.00035801532794721425
step: 410, loss: 1.4625260519096628e-05
step: 420, loss: 1.2345489267318044e-05
step: 430, loss: 6.181476783240214e-05
step: 440, loss: 3.9146285416791216e-05
step: 450, loss: 1.933016028488055e-05
step: 460, loss: 1.6957203115453012e-05
step: 470, loss: 4.6121196646708995e-05
step: 480, loss: 2.4812057745293714e-05
step: 490, loss: 3.1491461413679644e-05
step: 500, loss: 1.4476272554020397e-05
step: 510, loss: 2.3114096620702185e-05
step: 520, loss: 1.475567933084676e-05
step: 530, loss: 1.6357509593944997e-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=0.9294605809128631, f1=0.9247015610651974,
best_f1=0.9273897058823529
```


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

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

step: 0, loss: 3.0789757147431374e-05
step: 10, loss: 1.644689473323524e-05
step: 20, loss: 2.5582261514500715e-05
step: 30, loss: 2.764306373137515e-05
step: 40, loss: 0.030485982075333595
step: 50, loss: 2.9182287107687443e-05
step: 60, loss: 1.1578070370887872e-05
step: 70, loss: 0.00027608033269643784
step: 80, loss: 1.540383709652815e-05
step: 90, loss: 1.8562783225206658e-05
step: 100, loss: 2.767964542726986e-05
step: 110, loss: 0.00010680985724320635
step: 120, loss: 0.01481596753001213
step: 130, loss: 0.08010175824165344
step: 140, loss: 1.7039052181644365e-05
step: 150, loss: 2.425438469799701e-05
step: 160, loss: 1.8447190086590126e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 170, loss: 1.925920878420584e-05
step: 180, loss: 1.0736198419181164e-05
step: 190, loss: 2.8054162612534128e-05
step: 200, loss: 1.2926630915899295e-05
step: 210, loss: 4.7966652346076444e-05
step: 220, loss: 3.097073931712657e-05
step: 230, loss: 1.5016412362456322e-05
step: 240, loss: 1.9564529793569818e-05
step: 250, loss: 2.136726470780559e-05
step: 260, loss: 1.450237959943479e-05
step: 270, loss: 2.1162753910175525e-05
step: 280, loss: 1.840626100602094e-05
step: 290, loss: 1.2550359315355308e-05
step: 300, loss: 1.1943157915084157e-05
step: 310, loss: 1.6778449207777157e-05
step: 320, loss: 1.4822592675045598e-05
step: 330, loss: 1.77468427864369e-05
step: 340, loss: 1.1458888366178144e-05
step: 350, loss: 9.849588423094247e-06
step: 360, loss: 0.00012583770148921758
step: 370, loss: 1.5880308637861162e-05
step: 380, loss: 1.8156646547140554e-05
step: 390, loss: 3.95232746086549e-05
step: 400, loss: 2.5960012862924486e-05
step: 410, loss: 1.8935270418296568e-05
step: 420, loss: 1.6387282812502235e-05
step: 430, loss: 3.7188296118984e-05

```

step: 440, loss: 1.611165498616174e-05
step: 450, loss: 1.1451439604570623e-05
step: 460, loss: 1.6569811123190448e-05
step: 470, loss: 0.0014208215288817883
step: 480, loss: 1.3448105164570734e-05
step: 490, loss: 1.3310322174220346e-05
step: 500, loss: 1.980686283786781e-05
step: 510, loss: 1.5318077203119174e-05
step: 520, loss: 1.5854582670726813e-05
step: 530, loss: 1.5973833797033876e-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=0.9288354898336414, f1=0.920517560073937,
best_f1=0.9273897058823529
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, 196.76it/s]
load_f1 = 0.9290976058931861
real_f1 = 0.928801102434543

```

```
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:23, 189.89it/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: 0.5590924620628357
step: 10, loss: 0.3788471519947052
step: 20, loss: 0.35555291175842285
step: 30, loss: 0.3202214539051056
step: 40, loss: 0.1708986610174179
step: 50, loss: 0.4165574908256531
step: 60, loss: 0.27367472648620605
step: 70, loss: 0.15921321511268616
step: 80, loss: 0.19384491443634033
step: 90, loss: 0.4316141605377197
step: 100, loss: 0.3469602167606354
step: 110, loss: 0.19710111618041992
step: 120, loss: 0.2479216605424881
step: 130, loss: 0.16350336372852325
step: 140, loss: 0.2183479368686676
step: 150, loss: 0.20442163944244385
step: 160, loss: 0.40174299478530884
step: 170, loss: 0.264134019613266
step: 180, loss: 0.18820112943649292
step: 190, loss: 0.30141642689704895
step: 200, loss: 0.3057123124599457
step: 210, loss: 0.2319084256887436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5652173913043478, f1=0.5918367346938775,
best_f1=0.5918367346938775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1129695251584053
step: 10, loss: 0.21500715613365173
step: 20, loss: 0.13155905902385712
step: 30, loss: 0.24512116611003876
step: 40, loss: 0.1736481785774231
step: 50, loss: 0.21717245876789093
step: 60, loss: 0.3929520845413208
step: 70, loss: 0.1393406093120575
step: 80, loss: 0.15166085958480835
step: 90, loss: 0.06469780951738358
step: 100, loss: 0.03471573069691658
step: 110, loss: 0.09817763417959213
step: 120, loss: 0.17431136965751648
step: 130, loss: 0.030292386189103127
step: 140, loss: 0.11218669265508652
step: 150, loss: 0.282514363527298
step: 160, loss: 0.21046514809131622
step: 170, loss: 0.09330980479717255
step: 180, loss: 0.18660902976989746
step: 190, loss: 0.20271643996238708
step: 200, loss: 0.03335237875580788
step: 210, loss: 0.18149052560329437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5964912280701755, f1=0.6175869120654397,
best_f1=0.6175869120654397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06233687326312065
step: 10, loss: 0.1414567530155182
step: 20, loss: 0.1377442330121994

```

```

step: 30, loss: 0.11876001209020615
step: 40, loss: 0.04959765076637268
step: 50, loss: 0.0649949312210083
step: 60, loss: 0.14186246693134308
step: 70, loss: 0.13395054638385773
step: 80, loss: 0.10470361262559891
step: 90, loss: 0.04131290316581726
step: 100, loss: 0.23510627448558807
step: 110, loss: 0.1234106793999672
step: 120, loss: 0.06867152452468872
step: 130, loss: 0.08159685879945755
step: 140, loss: 0.06724978238344193
step: 150, loss: 0.16192883253097534
step: 160, loss: 0.02506379410624504
step: 170, loss: 0.10674363374710083
step: 180, loss: 0.13503296673297882
step: 190, loss: 0.1804877370595932
step: 200, loss: 0.03337739780545235
step: 210, loss: 0.12475192546844482
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6048951048951048, f1=0.5924596050269301,
best_f1=0.5924596050269301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07037438452243805
step: 10, loss: 0.034259941428899765
step: 20, loss: 0.027360081672668457
step: 30, loss: 0.07753288745880127
step: 40, loss: 0.014488773420453072
step: 50, loss: 0.1619146764278412
step: 60, loss: 0.17716985940933228
step: 70, loss: 0.2243281453847885
step: 80, loss: 0.049238912761211395
step: 90, loss: 0.012851041741669178
step: 100, loss: 0.10581080615520477
step: 110, loss: 0.15385296940803528
step: 120, loss: 0.11151090264320374
step: 130, loss: 0.1520218551158905
step: 140, loss: 0.08734625577926636
step: 150, loss: 0.055153556168079376
step: 160, loss: 0.04917468503117561

```

```

step: 170, loss: 0.11825382709503174
step: 180, loss: 0.28122377395629883
step: 190, loss: 0.09033510088920593
step: 200, loss: 0.18335723876953125
step: 210, loss: 0.09388848394155502
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6317757009345795, f1=0.622568093385214,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12799789011478424
step: 10, loss: 0.06078709661960602
step: 20, loss: 0.1661740243434906
step: 30, loss: 0.026084640994668007
step: 40, loss: 0.07323455065488815
step: 50, loss: 0.06115838885307312
step: 60, loss: 0.10030581802129745
step: 70, loss: 0.03887937217950821
step: 80, loss: 0.1170053482055664
step: 90, loss: 0.08971276134252548
step: 100, loss: 0.01165288221091032
step: 110, loss: 0.11820022016763687
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.140786811709404
step: 130, loss: 0.15925461053848267
step: 140, loss: 0.12074024230241776
step: 150, loss: 0.0527220219373703
step: 160, loss: 0.16714462637901306
step: 170, loss: 0.15420173108577728
step: 180, loss: 0.0531000941991806
step: 190, loss: 0.05363060161471367
step: 200, loss: 0.05666341632604599
step: 210, loss: 0.029221754521131516
/usr/local/lib/python3.7/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=0.6161971830985915, f1=0.5693950177935942,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.04859522357583046
step: 10, loss: 0.03887087479233742
step: 20, loss: 0.06923296302556992
step: 30, loss: 0.0025536406319588423
step: 40, loss: 0.013050488196313381
step: 50, loss: 0.009347464889287949
step: 60, loss: 0.1984003335237503
step: 70, loss: 0.04436331242322922
step: 80, loss: 0.035453855991363525
step: 90, loss: 0.2120271772146225
step: 100, loss: 0.021692121401429176
step: 110, loss: 0.012622465379536152
step: 120, loss: 0.037720777094364166
step: 130, loss: 0.07641693949699402
step: 140, loss: 0.041147150099277496
step: 150, loss: 0.016373103484511375
step: 160, loss: 0.01723695732653141
step: 170, loss: 0.10900861024856567
step: 180, loss: 0.010970264673233032
step: 190, loss: 0.02787916734814644
step: 200, loss: 0.02456754818558693
step: 210, loss: 0.009517267346382141
/usr/local/lib/python3.7/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=0.5979797979797981, f1=0.6012269938650306,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002802124712616205
step: 10, loss: 0.004655301570892334
step: 20, loss: 0.030624235048890114
step: 30, loss: 0.013798044994473457
step: 40, loss: 0.0008616087725386024
step: 50, loss: 0.1795848309993744
step: 60, loss: 0.09272187948226929
step: 70, loss: 0.01455574482679367
step: 80, loss: 0.029834512621164322
step: 90, loss: 0.08699300140142441
step: 100, loss: 0.02471756935119629
step: 110, loss: 0.1683957278728485
step: 120, loss: 0.060408953577280045
step: 130, loss: 0.012270599603652954
step: 140, loss: 0.0573556125164032
step: 150, loss: 0.01705934852361679
step: 160, loss: 0.025066424161195755
step: 170, loss: 0.0012401192216202617

```



```

step: 180, loss: 0.009936521761119366
step: 190, loss: 0.20835043489933014
step: 200, loss: 0.022121695801615715
step: 210, loss: 0.013226390816271305
/usr/local/lib/python3.7/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=0.6046511627906976, f1=0.6090373280943024,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003066035220399499
step: 10, loss: 0.036620207130908966
step: 20, loss: 0.003569790394976735
step: 30, loss: 0.002728547900915146
step: 40, loss: 0.0033121779561042786
step: 50, loss: 0.012561982497572899
step: 60, loss: 0.002226373879238963
step: 70, loss: 0.01549326628446579
step: 80, loss: 0.021660545840859413
step: 90, loss: 0.008796717040240765
step: 100, loss: 0.04575921967625618
step: 110, loss: 0.11028692126274109
step: 120, loss: 0.017255134880542755
step: 130, loss: 0.0018712584860622883
step: 140, loss: 0.04853309690952301
step: 150, loss: 0.007341315969824791
step: 160, loss: 0.052344854921102524
step: 170, loss: 0.020526684820652008
step: 180, loss: 0.09134487807750702
step: 190, loss: 0.005968289449810982
step: 200, loss: 0.0029266367200762033
step: 210, loss: 0.08823520690202713
/usr/local/lib/python3.7/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=0.5902335456475584, f1=0.5974025974025974,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05991426482796669
step: 10, loss: 0.02178001031279564
step: 20, loss: 0.0034706543665379286
step: 30, loss: 0.03634351119399071
step: 40, loss: 0.018322410061955452
step: 50, loss: 0.004910433664917946

```

```

step: 60, loss: 0.02252560295164585
step: 70, loss: 0.0038726578932255507
step: 80, loss: 0.001836448791436851
step: 90, loss: 0.000957471551373601
step: 100, loss: 0.06056511402130127
step: 110, loss: 0.019644295796751976
step: 120, loss: 0.007551337592303753
step: 130, loss: 0.059359192848205566
step: 140, loss: 0.021715031936764717
step: 150, loss: 0.12270450592041016
step: 160, loss: 0.004067180212587118
step: 170, loss: 0.0007677412359043956
step: 180, loss: 0.05218027904629707
step: 190, loss: 0.0017936421791091561
step: 200, loss: 0.010583673603832722
step: 210, loss: 0.0005627088248729706
/usr/local/lib/python3.7/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=0.5842217484008528, f1=0.5986984815618221,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005959575646556914
step: 10, loss: 0.0024538652505725622
step: 20, loss: 0.00031818568822927773
step: 30, loss: 0.006532663945108652
step: 40, loss: 0.0005341543001122773
step: 50, loss: 0.020652366802096367
step: 60, loss: 0.00567589420825243
step: 70, loss: 0.03681771457195282
step: 80, loss: 0.0018455700483173132
step: 90, loss: 0.01307629980146885
step: 100, loss: 0.02228493057191372
step: 110, loss: 0.0003114108694717288
step: 120, loss: 0.00033975334372371435
step: 130, loss: 0.008436412550508976
step: 140, loss: 0.002812353428453207
step: 150, loss: 0.004715615417808294
step: 160, loss: 0.003038896946236491
step: 170, loss: 0.017969565466046333
step: 180, loss: 0.02635512501001358
step: 190, loss: 0.06967788189649582
step: 200, loss: 0.0023397470358759165
step: 210, loss: 0.029781492426991463
/usr/local/lib/python3.7/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=0.5831622176591376, f1=0.6068376068376068,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028371257707476616
step: 10, loss: 0.02738381177186966
step: 20, loss: 0.00878753513097763
step: 30, loss: 0.0004606659640558064
step: 40, loss: 0.018737919628620148
step: 50, loss: 0.0012604553485289216
step: 60, loss: 0.0027703389059752226
step: 70, loss: 0.002990937326103449
step: 80, loss: 0.03574798256158829
step: 90, loss: 0.0002099749690387398
step: 100, loss: 0.011633596383035183
step: 110, loss: 0.0059507545083761215
step: 120, loss: 0.027009882032871246
step: 130, loss: 0.0019039887702092528
step: 140, loss: 0.00015032815281301737
step: 150, loss: 0.00019315083045512438
step: 160, loss: 0.0007277224212884903
step: 170, loss: 0.0005207165377214551
step: 180, loss: 0.0043829018250107765
step: 190, loss: 0.00196390924975276
step: 200, loss: 0.036338187754154205
step: 210, loss: 0.0014137544203549623
/usr/local/lib/python3.7/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=0.604914933837429, f1=0.6039215686274509,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020389100536704063
step: 10, loss: 0.001741307321935892
step: 20, loss: 0.017858335748314857
step: 30, loss: 0.02418668568134308
step: 40, loss: 0.0011570433853194118
step: 50, loss: 0.022174960002303123
step: 60, loss: 0.0005898622912354767
step: 70, loss: 0.0019709060434252024
step: 80, loss: 0.0008519911207258701
step: 90, loss: 0.021457916125655174
step: 100, loss: 0.22634465992450714
step: 110, loss: 0.004880909807980061

```

```

step: 120, loss: 0.0028822312597185373
step: 130, loss: 0.00669308053329587
step: 140, loss: 0.0002459526585880667
step: 150, loss: 0.0012323344126343727
step: 160, loss: 0.028545280918478966
step: 170, loss: 0.0007660348201170564
step: 180, loss: 0.0010789104271680117
step: 190, loss: 0.04456828907132149
step: 200, loss: 0.0030412902124226093
step: 210, loss: 0.005509525071829557
/usr/local/lib/python3.7/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=0.5991561181434599, f1=0.6004228329809724,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018490031361579895
step: 10, loss: 0.00014806863327976316
step: 20, loss: 0.011868514120578766
step: 30, loss: 0.0014819809002801776
step: 40, loss: 0.031928323209285736
step: 50, loss: 0.0018534479895606637
step: 60, loss: 0.001031358609907329
step: 70, loss: 0.06434672325849533
step: 80, loss: 0.0028558073099702597
step: 90, loss: 0.0001467307738494128
step: 100, loss: 0.00022561968944501132
step: 110, loss: 0.0010408916277810931
step: 120, loss: 0.0007470757700502872
step: 130, loss: 0.00043596018804237247
step: 140, loss: 0.00033667802927084267
step: 150, loss: 0.0030599385499954224
step: 160, loss: 0.04745488613843918
step: 170, loss: 0.0017310489201918244
step: 180, loss: 0.012893936596810818
step: 190, loss: 0.0003572539135348052
step: 200, loss: 0.0008229516679421067
step: 210, loss: 0.0010084344539791346
/usr/local/lib/python3.7/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=0.5887640449438203, f1=0.5947136563876653,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 9.978132584365085e-05
step: 10, loss: 0.003999378066509962
step: 20, loss: 0.0002165373443858698
step: 30, loss: 0.0004922226071357727
step: 40, loss: 0.00045453401980921626
step: 50, loss: 0.00039587978972122073
step: 60, loss: 0.004506767261773348
step: 70, loss: 0.00033577598514966667
step: 80, loss: 0.00033415385405533016
step: 90, loss: 0.0035019880160689354
step: 100, loss: 0.0011144704185426235
step: 110, loss: 0.001425613765604794
step: 120, loss: 0.03990292549133301
step: 130, loss: 0.0014697350561618805
step: 140, loss: 0.0008285561925731599
step: 150, loss: 0.028536800295114517
step: 160, loss: 0.004388016182929277
step: 170, loss: 0.000502069597132504
step: 180, loss: 0.0003402537840884179
step: 190, loss: 0.030032198876142502
step: 200, loss: 0.007054335903376341
step: 210, loss: 0.010737268254160881
/usr/local/lib/python3.7/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=0.5936842105263158, f1=0.6083333333333334,
best_f1=0.622568093385214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033286598045378923
step: 10, loss: 0.00018632158753462136
step: 20, loss: 0.0008595044491812587
step: 30, loss: 0.020967302843928337
step: 40, loss: 0.00031627604039385915
step: 50, loss: 0.00012316094944253564
step: 60, loss: 0.003360751550644636
step: 70, loss: 0.0003249008732382208
step: 80, loss: 0.0006368382018990815
step: 90, loss: 0.0027672662399709225
step: 100, loss: 0.00022562764934264123
step: 110, loss: 0.00013094086898490787
step: 120, loss: 0.0001719384454190731
step: 130, loss: 0.00039110175566747785
step: 140, loss: 0.00025869658566080034
step: 150, loss: 0.00015413929941132665
step: 160, loss: 0.00025858331355266273
step: 170, loss: 0.0002446213911753148

```

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.18809722363948822
step: 190, loss: 0.00013197652879171073
step: 200, loss: 0.007127661257982254
step: 210, loss: 0.0006931893294677138
/usr/local/lib/python3.7/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=0.5917926565874729, f1=0.5970149253731343,
best_f1=0.622568093385214
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, 282.12it/s]
load_f1 = 0.62708719851577
real_f1 = 0.6303939962476548
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.72it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6038344502449036
step: 10, loss: 0.35017532110214233
step: 20, loss: 0.2715592682361603
step: 30, loss: 0.45775747299194336
step: 40, loss: 0.42936789989471436
step: 50, loss: 0.3434927761554718
step: 60, loss: 0.2798023521900177
step: 70, loss: 0.2740243375301361
step: 80, loss: 0.22001154720783234
step: 90, loss: 0.27973175048828125
step: 100, loss: 0.317672461271286
step: 110, loss: 0.35004350543022156
step: 120, loss: 0.17725518345832825
step: 130, loss: 0.15874965488910675
step: 140, loss: 0.06637181341648102
step: 150, loss: 0.12555311620235443
step: 160, loss: 0.12995010614395142
step: 170, loss: 0.23692256212234497
step: 180, loss: 0.03440112993121147
step: 190, loss: 0.30314207077026367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6700767263427111, f1=0.696103896103896,
best_f1=0.696103896103896

```



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

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

```
step: 0, loss: 0.2866505980491638  
step: 10, loss: 0.06993915885686874  
step: 20, loss: 0.1366489976644516  
step: 30, loss: 0.18200545012950897  
step: 40, loss: 0.16216707229614258  
step: 50, loss: 0.05553894490003586  
step: 60, loss: 0.24920590221881866  
step: 70, loss: 0.16984480619430542  
step: 80, loss: 0.17196887731552124  
step: 90, loss: 0.1540011465549469  
step: 100, loss: 0.039440326392650604  
step: 110, loss: 0.19511893391609192  
step: 120, loss: 0.24440671503543854  
step: 130, loss: 0.12939487397670746  
step: 140, loss: 0.07291470468044281  
step: 150, loss: 0.09700916707515717  
step: 160, loss: 0.10869219154119492  
step: 170, loss: 0.060631632804870605  
step: 180, loss: 0.1954669952392578  
step: 190, loss: 0.05332573503255844
```

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

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

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

```
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=0.7570332480818414, f1=0.7616580310880829,  
best_f1=0.7616580310880829
```

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

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

```
step: 0, loss: 0.05574587732553482  
step: 10, loss: 0.1839180439710617  
step: 20, loss: 0.13458497822284698  
step: 30, loss: 0.051852863281965256  
step: 40, loss: 0.1338832527399063  
step: 50, loss: 0.1448737382888794  
step: 60, loss: 0.032624926418066025  
step: 70, loss: 0.05789513513445854  
step: 80, loss: 0.144624263048172  
step: 90, loss: 0.10477012395858765  
step: 100, loss: 0.01987914741039276  
step: 110, loss: 0.01346411369740963  
step: 120, loss: 0.021878071129322052
```

```

step: 130, loss: 0.013012605719268322
step: 140, loss: 0.01139745395630598
step: 150, loss: 0.16934578120708466
step: 160, loss: 0.08496378362178802
step: 170, loss: 0.21114617586135864
step: 180, loss: 0.06802897900342941
step: 190, loss: 0.08922412246465683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7788944723618091, f1=0.781725888324873,
best_f1=0.781725888324873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05951181426644325
step: 10, loss: 0.0826515257358551
step: 20, loss: 0.12221337109804153
step: 30, loss: 0.040649283677339554
step: 40, loss: 0.030574627220630646
step: 50, loss: 0.1413862556219101
step: 60, loss: 0.11799022555351257
step: 70, loss: 0.03825432062149048
step: 80, loss: 0.06849279999732971
step: 90, loss: 0.01529902033507824
step: 100, loss: 0.008093807846307755
step: 110, loss: 0.004545166622847319
step: 120, loss: 0.031082328408956528
step: 130, loss: 0.07338318973779678
step: 140, loss: 0.02394293248653412
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.04061061143875122
step: 160, loss: 0.034691616892814636
step: 170, loss: 0.14576487243175507
step: 180, loss: 0.1859397143125534
step: 190, loss: 0.10667972266674042
/usr/local/lib/python3.7/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=0.7614213197969544, f1=0.7696335078534032,
best_f1=0.781725888324873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.05628993362188339
step: 10, loss: 0.013861773535609245
step: 20, loss: 0.0395953506231308
step: 30, loss: 0.016098838299512863
step: 40, loss: 0.07240705192089081
step: 50, loss: 0.1980750411748886
step: 60, loss: 0.016094839200377464
step: 70, loss: 0.006672574672847986
step: 80, loss: 0.011747248470783234
step: 90, loss: 0.004169085528701544
step: 100, loss: 0.00250050937756896
step: 110, loss: 0.0071774981915950775
step: 120, loss: 0.005610847380012274
step: 130, loss: 0.1130940243601799
step: 140, loss: 0.028937002643942833
step: 150, loss: 0.013678775168955326
step: 160, loss: 0.1018708273768425
step: 170, loss: 0.0019111447036266327
step: 180, loss: 0.013929974287748337
step: 190, loss: 0.08244487643241882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8097826086956521, f1=0.7978436657681941,
best_f1=0.7978436657681941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0156876090914011
step: 10, loss: 0.17347535490989685
step: 20, loss: 0.00862239208072424
step: 30, loss: 0.0008386505651287735
step: 40, loss: 0.07178957015275955
step: 50, loss: 0.022650932893157005
step: 60, loss: 0.01103738322854042
step: 70, loss: 0.06556116789579391
step: 80, loss: 0.0356634221971035
step: 90, loss: 0.007329197134822607
step: 100, loss: 0.0016532412264496088
step: 110, loss: 0.013296172954142094
step: 120, loss: 0.0043404605239629745
step: 130, loss: 0.00849572941660881
step: 140, loss: 0.01977849379181862
step: 150, loss: 0.0012131640687584877

```

```

step: 160, loss: 0.051714953035116196
step: 170, loss: 0.02540382742881775
step: 180, loss: 0.006146129686385393
step: 190, loss: 0.034208428114652634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8142076502732241, f1=0.7722222222222223,
best_f1=0.7722222222222223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004687666427344084
step: 10, loss: 0.05074580758810043
step: 20, loss: 0.09888456761837006
step: 30, loss: 0.0013615497155115008
step: 40, loss: 0.08408064395189285
step: 50, loss: 0.006513007916510105
step: 60, loss: 0.027552573010325432
step: 70, loss: 0.00419675512239337
step: 80, loss: 0.006944779772311449
step: 90, loss: 0.007735304068773985
step: 100, loss: 0.0016948094125837088
step: 110, loss: 0.010794227942824364
step: 120, loss: 0.0010750622022897005
step: 130, loss: 0.00895669311285019
step: 140, loss: 0.014861628413200378
step: 150, loss: 0.018635345622897148
step: 160, loss: 0.023238398134708405
step: 170, loss: 0.003018451388925314
step: 180, loss: 0.0009943261975422502
step: 190, loss: 0.0009419245179742575
/usr/local/lib/python3.7/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=0.7977528089887641, f1=0.793388429752066,
best_f1=0.7722222222222223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022958205081522465
step: 10, loss: 0.005293913185596466
step: 20, loss: 0.0014985427260398865
step: 30, loss: 0.007070038467645645

```

```

step: 40, loss: 0.0008490215986967087
step: 50, loss: 0.002504267729818821
step: 60, loss: 0.0017501991242170334
step: 70, loss: 0.06039140745997429
step: 80, loss: 0.00044610502663999796
step: 90, loss: 0.006708289030939341
step: 100, loss: 0.0014727701200172305
step: 110, loss: 0.0005374825559556484
step: 120, loss: 0.0016992503078654408
step: 130, loss: 0.0004411455010995269
step: 140, loss: 0.001318500842899084
step: 150, loss: 0.09237765520811081
step: 160, loss: 0.0007494536694139242
step: 170, loss: 0.0012380536645650864
step: 180, loss: 0.04594253748655319
step: 190, loss: 0.07998346537351608
/usr/local/lib/python3.7/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=0.7851458885941645, f1=0.7712765957446809,
best_f1=0.7722222222222223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003315781708806753
step: 10, loss: 0.023590462282299995
step: 20, loss: 0.0012184525839984417
step: 30, loss: 0.00043794390512630343
step: 40, loss: 0.0009224951500073075
step: 50, loss: 0.0010112002491950989
step: 60, loss: 0.0014340955531224608
step: 70, loss: 0.0006675987388007343
step: 80, loss: 0.0007677605026401579
step: 90, loss: 0.07965237647294998
step: 100, loss: 0.015889819711446762
step: 110, loss: 0.0010777920251712203
step: 120, loss: 0.0015872593503445387
step: 130, loss: 0.0009206653339788318
step: 140, loss: 0.008350096642971039
step: 150, loss: 0.0005297409952618182
step: 160, loss: 0.00033751249429769814
step: 170, loss: 0.0010871633421629667
step: 180, loss: 0.0012494507245719433
step: 190, loss: 0.0006729209562763572
/usr/local/lib/python3.7/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=0.8072916666666666, f1=0.7855297157622738,

```

```

best_f1=0.7722222222222223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004184326680842787
step: 10, loss: 0.0016902424395084381
step: 20, loss: 0.0031089133117347956
step: 30, loss: 0.00022060533228795975
step: 40, loss: 0.0025761539582163095
step: 50, loss: 0.0003890780499204993
step: 60, loss: 0.0003693464968819171
step: 70, loss: 0.0012539072195068002
step: 80, loss: 0.0013376668794080615
step: 90, loss: 0.013101840391755104
step: 100, loss: 0.00047198584070429206
step: 110, loss: 0.000991112319752574
step: 120, loss: 0.09727384150028229
step: 130, loss: 0.00038899757782928646
step: 140, loss: 0.0004054451419506222
step: 150, loss: 0.1711529642343521
step: 160, loss: 0.002984993625432253
step: 170, loss: 0.0264178104698658
step: 180, loss: 0.0004983368562534451
step: 190, loss: 0.000552344077732414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.827027027027027, f1=0.7999999999999999,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003705908020492643
step: 10, loss: 0.0012042816961184144
step: 20, loss: 0.004894117824733257
step: 30, loss: 0.005296419840306044
step: 40, loss: 0.00021445816673804075
step: 50, loss: 0.015013475902378559
step: 60, loss: 0.00026035262271761894
step: 70, loss: 0.0012837356189265847
step: 80, loss: 0.0004498417256399989
step: 90, loss: 0.0003372995415702462
step: 100, loss: 0.00035705018672160804
step: 110, loss: 0.000279326835880056

```

```

step: 120, loss: 0.0004636045196093619
step: 130, loss: 0.0003609077539294958
step: 140, loss: 0.0002742950455285609
step: 150, loss: 0.00028726071468554437
step: 160, loss: 0.00035407079849392176
step: 170, loss: 0.00023450484150089324
step: 180, loss: 0.007516215555369854
step: 190, loss: 0.0003287677827756852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8283378746594007, f1=0.7978436657681941,
best_f1=0.7978436657681941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011096570233348757
step: 10, loss: 0.00045436384971253574
step: 20, loss: 0.00014550946070812643
step: 30, loss: 0.00022780393192078918
step: 40, loss: 0.00025090196868404746
step: 50, loss: 0.0004000073822680861
step: 60, loss: 0.00030608067754656076
step: 70, loss: 0.00016314905951730907
step: 80, loss: 0.003022619755938649
step: 90, loss: 0.0006568431854248047
step: 100, loss: 0.0002699411998037249
step: 110, loss: 0.0021726808045059443
step: 120, loss: 0.0003946149372495711
step: 130, loss: 0.0007218809332698584
step: 140, loss: 0.004055683966726065
step: 150, loss: 0.0003714847262017429
step: 160, loss: 0.00012358093226794153
step: 170, loss: 0.0015653324080631137
step: 180, loss: 0.00011071354674641043
step: 190, loss: 0.0002455060603097081
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8074866310160429, f1=0.8064516129032259,
best_f1=0.7978436657681941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0014517382951453328
step: 10, loss: 0.00033493011142127216
step: 20, loss: 0.0003480854502413422
step: 30, loss: 0.0009037598501890898
step: 40, loss: 0.0030680804047733545
step: 50, loss: 0.0003650157304946333
step: 60, loss: 0.00019236010848544538
step: 70, loss: 0.005168217699974775
step: 80, loss: 0.002211923711001873
step: 90, loss: 0.00041171431075781584
step: 100, loss: 0.0009057394927367568
step: 110, loss: 0.00014765870582778007
step: 120, loss: 0.0021073929965496063
step: 130, loss: 0.00010746010957518592
step: 140, loss: 0.0009295449126511812
step: 150, loss: 0.00014540596748702228
step: 160, loss: 0.014327836222946644
step: 170, loss: 0.00038273577229119837
step: 180, loss: 0.00010961421503452584
step: 190, loss: 0.006107485853135586
/usr/local/lib/python3.7/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=0.8111111111111112, f1=0.7932960893854748,
best_f1=0.7978436657681941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000506409618537873
step: 10, loss: 0.00012024952593492344
step: 20, loss: 0.00020144552399870008
step: 30, loss: 0.00011811555305030197
step: 40, loss: 0.001144975540228188
step: 50, loss: 0.00037750022602267563
step: 60, loss: 0.0002546116302255541
step: 70, loss: 0.00034126153332181275
step: 80, loss: 0.00030485709430649877
step: 90, loss: 0.00111105025280267
step: 100, loss: 0.0078098708763718605
step: 110, loss: 0.0002484606229700148
step: 120, loss: 0.0002285455702804029
step: 130, loss: 0.008626549504697323
step: 140, loss: 0.000326736131682992
step: 150, loss: 0.00033519690623506904
step: 160, loss: 0.0015948378713801503
step: 170, loss: 0.0002657890727277845
step: 180, loss: 0.0024104544427245855
step: 190, loss: 0.00023224359028972685

```



```

/usr/local/lib/python3.7/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=0.8162162162162163, f1=0.7988980716253442,
best_f1=0.7978436657681941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012296802597120404
step: 10, loss: 0.00022616132628172636
step: 20, loss: 0.00032929808367043734
step: 30, loss: 0.00012574854190461338
step: 40, loss: 9.556335862725973e-05
step: 50, loss: 0.00030527188209816813
step: 60, loss: 0.000506063865032047
step: 70, loss: 0.01630699448287487
step: 80, loss: 0.00025547383120283484
step: 90, loss: 0.00010894666775129735
step: 100, loss: 0.0007596118957735598
step: 110, loss: 0.00016577982751186937
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 9.875404794001952e-05
step: 130, loss: 0.00024688043049536645
step: 140, loss: 0.002163863042369485
step: 150, loss: 0.00040724503924138844
step: 160, loss: 0.00019300970598123968
step: 170, loss: 0.0004650020564440638
step: 180, loss: 0.0019756571855396032
step: 190, loss: 0.00012260272342246026
/usr/local/lib/python3.7/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=0.8133704735376045, f1=0.7910863509749304,
best_f1=0.7978436657681941
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 164.43it/s]
load_f1 = 0.8264462809917356
real_f1 = 0.8232044198895028
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 186.43it/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: 0.6169766187667847

```

step: 10, loss: 0.3609139621257782
step: 20, loss: 0.2963274121284485
step: 30, loss: 0.36663150787353516
step: 40, loss: 0.26752370595932007
step: 50, loss: 0.24720355868339539
step: 60, loss: 0.2527465224266052
step: 70, loss: 0.35222703218460083
step: 80, loss: 0.37202006578445435
step: 90, loss: 0.2159721702337265
step: 100, loss: 0.19585421681404114
step: 110, loss: 0.22838014364242554
step: 120, loss: 0.12382663786411285
step: 130, loss: 0.08099700510501862
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.23456795513629913
step: 150, loss: 0.25511595606803894
step: 160, loss: 0.20600946247577667
step: 170, loss: 0.14765296876430511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7072599531615924, f1=0.7294117647058824,
best_f1=0.7294117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08232636004686356
step: 10, loss: 0.21141158044338226
step: 20, loss: 0.034188784658908844
step: 30, loss: 0.19880151748657227
step: 40, loss: 0.11937592178583145
step: 50, loss: 0.17250648140907288
step: 60, loss: 0.05537216737866402
step: 70, loss: 0.1524004191160202
step: 80, loss: 0.04225882515311241
step: 90, loss: 0.09858202934265137
step: 100, loss: 0.1250404715538025
step: 110, loss: 0.04151167348027229
step: 120, loss: 0.08941173553466797
step: 130, loss: 0.05915664881467819
step: 140, loss: 0.29104286432266235
step: 150, loss: 0.23402152955532074
step: 160, loss: 0.11328577250242233
step: 170, loss: 0.04569673910737038

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7684478371501272, f1=0.7641025641025642,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07455157488584518
step: 10, loss: 0.03568779304623604
step: 20, loss: 0.02094150520861149
step: 30, loss: 0.19865120947360992
step: 40, loss: 0.08782533556222916
step: 50, loss: 0.08339294046163559
step: 60, loss: 0.13512353599071503
step: 70, loss: 0.07756906002759933
step: 80, loss: 0.053874872624874115
step: 90, loss: 0.08863933384418488
step: 100, loss: 0.009021107107400894
step: 110, loss: 0.11679885536432266
step: 120, loss: 0.06308195739984512
step: 130, loss: 0.09288956969976425
step: 140, loss: 0.13570202887058258
step: 150, loss: 0.02733656018972397
step: 160, loss: 0.015585382468998432
step: 170, loss: 0.06934770196676254
/usr/local/lib/python3.7/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=0.7405405405405405, f1=0.7745358090185676,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010014748200774193
step: 10, loss: 0.08738133311271667
step: 20, loss: 0.013822992332279682
step: 30, loss: 0.04675177484750748
step: 40, loss: 0.0021393944043666124
step: 50, loss: 0.26736506819725037
step: 60, loss: 0.1621464639902115
step: 70, loss: 0.004816029220819473
step: 80, loss: 0.025243734940886497
step: 90, loss: 0.08554710447788239

```

```

step: 100, loss: 0.12797288596630096
step: 110, loss: 0.12136855721473694
step: 120, loss: 0.010613437741994858
step: 130, loss: 0.1364338994026184
step: 140, loss: 0.027881663292646408
step: 150, loss: 0.08244501054286957
step: 160, loss: 0.24660897254943848
step: 170, loss: 0.014413172379136086
/usr/local/lib/python3.7/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=0.7524752475247525, f1=0.7817745803357313,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014362513087689877
step: 10, loss: 0.01141625177115202
step: 20, loss: 0.012127174064517021
step: 30, loss: 0.15125772356987
step: 40, loss: 0.03064757026731968
step: 50, loss: 0.03945663571357727
step: 60, loss: 0.02619469352066517
step: 70, loss: 0.25246331095695496
step: 80, loss: 0.11597655713558197
step: 90, loss: 0.1040320098400116
step: 100, loss: 0.01807698979973793
step: 110, loss: 0.06973173469305038
step: 120, loss: 0.010964103043079376
step: 130, loss: 0.0686064139008522
step: 140, loss: 0.00676066754385829
step: 150, loss: 0.013936156406998634
step: 160, loss: 0.10828331112861633
step: 170, loss: 0.008933210745453835
/usr/local/lib/python3.7/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=0.748051948051948, f1=0.7969151670951157,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006159698124974966
step: 10, loss: 0.009616849012672901
step: 20, loss: 0.008465137332677841
step: 30, loss: 0.034556951373815536
step: 40, loss: 0.02265981025993824
step: 50, loss: 0.0355074405670166

```

```

step: 60, loss: 0.002760391216725111
step: 70, loss: 0.08572655916213989
step: 80, loss: 0.021818581968545914
step: 90, loss: 0.07362550497055054
step: 100, loss: 0.005717556923627853
step: 110, loss: 0.0016719831619411707
step: 120, loss: 0.005282358266413212
step: 130, loss: 0.007280683610588312
step: 140, loss: 0.028033411130309105
step: 150, loss: 0.03907712921500206
step: 160, loss: 0.06280979514122009
step: 170, loss: 0.0028916066512465477
/usr/local/lib/python3.7/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=0.7549999999999999, f1=0.7536945812807883,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000819300243165344
step: 10, loss: 0.0012687182752415538
step: 20, loss: 0.009683684445917606
step: 30, loss: 0.015627404674887657
step: 40, loss: 0.003998574800789356
step: 50, loss: 0.07463137805461884
step: 60, loss: 0.002249309793114662
step: 70, loss: 0.009632511995732784
step: 80, loss: 0.027839437127113342
step: 90, loss: 0.0002612208481878042
step: 100, loss: 0.008359747938811779
step: 110, loss: 0.023699507117271423
step: 120, loss: 0.005811577662825584
step: 130, loss: 0.10290469229221344
step: 140, loss: 0.0042922357097268105
step: 150, loss: 0.014056673273444176
step: 160, loss: 0.000561302702408284
step: 170, loss: 0.17600810527801514
/usr/local/lib/python3.7/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=0.7386934673366835, f1=0.7889447236180904,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007049474399536848
step: 10, loss: 0.009763794019818306

```

```

step: 20, loss: 0.00038532837061211467
step: 30, loss: 0.013582250103354454
step: 40, loss: 0.000492663006298244
step: 50, loss: 0.006989933550357819
step: 60, loss: 0.0010960528161376715
step: 70, loss: 0.0033255810849368572
step: 80, loss: 0.014244125224649906
step: 90, loss: 0.004634442273527384
step: 100, loss: 0.04138204827904701
step: 110, loss: 0.03242680802941322
step: 120, loss: 0.0004872340359725058
step: 130, loss: 0.000776361848693341
step: 140, loss: 0.0004890849813818932
step: 150, loss: 0.06716973334550858
step: 160, loss: 0.001014763256534934
step: 170, loss: 0.0020617626141756773
/usr/local/lib/python3.7/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=0.7530562347188264, f1=0.7819905213270142,
best_f1=0.7641025641025642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007940666982904077
step: 10, loss: 0.029993589967489243
step: 20, loss: 0.0032508561853319407
step: 30, loss: 0.022845886647701263
step: 40, loss: 0.006875941995531321
step: 50, loss: 0.0014149354537948966
step: 60, loss: 0.10964754223823547
step: 70, loss: 0.05197994410991669
step: 80, loss: 0.0012384160654619336
step: 90, loss: 0.02603832259774208
step: 100, loss: 0.0060099209658801556
step: 110, loss: 0.015001441352069378
step: 120, loss: 0.001994151622056961
step: 130, loss: 0.35352230072021484
step: 140, loss: 0.0003944489872083068
step: 150, loss: 0.018034713342785835
step: 160, loss: 0.17346368730068207
step: 170, loss: 0.004010668024420738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7792915531335151, f1=0.7957559681697612,
best_f1=0.7957559681697612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05896247923374176
step: 10, loss: 0.00838873628526926
step: 20, loss: 0.09144581109285355
step: 30, loss: 0.05650656670331955
step: 40, loss: 0.000889665971044451
step: 50, loss: 0.05006851255893707
step: 60, loss: 0.0007450677803717554
step: 70, loss: 0.0173883568495512
step: 80, loss: 0.001098880311474204
step: 90, loss: 0.01371186226606369
step: 100, loss: 0.0002209121303167194
step: 110, loss: 0.001054995460435748
step: 120, loss: 0.0006215343601070344
step: 130, loss: 0.004401498474180698
step: 140, loss: 0.03148452192544937
step: 150, loss: 0.01806139014661312
step: 160, loss: 0.004724104888737202
step: 170, loss: 0.001209661248140037
/usr/local/lib/python3.7/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=0.7493670886075949, f1=0.7791563275434243,
best_f1=0.7957559681697612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02667255699634552
step: 10, loss: 0.03189006820321083
step: 20, loss: 0.0026150611229240894
step: 30, loss: 0.00019661210535559803
step: 40, loss: 0.00022266226005740464
step: 50, loss: 0.011319456622004509
step: 60, loss: 0.02306235209107399
step: 70, loss: 0.00013700798444915563
step: 80, loss: 0.0004388083762023598
step: 90, loss: 0.0001511702430434525
step: 100, loss: 0.00026523828273639083
step: 110, loss: 0.025249967351555824
step: 120, loss: 0.00017823318194132298
step: 130, loss: 0.00015910889487713575
step: 140, loss: 0.00026642915327101946
step: 150, loss: 0.007523934356868267

```

```

step: 160, loss: 0.004268669057637453
step: 170, loss: 0.0010571142192929983
/usr/local/lib/python3.7/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=0.7463414634146343, f1=0.7684964200477328,
best_f1=0.7957559681697612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002678896998986602
step: 10, loss: 0.009116039611399174
step: 20, loss: 0.005588125437498093
step: 30, loss: 0.0005626385100185871
step: 40, loss: 0.0027874454390257597
step: 50, loss: 0.00018570994143374264
step: 60, loss: 0.0005580022698268294
step: 70, loss: 0.04348693788051605
step: 80, loss: 0.0001644780277274549
step: 90, loss: 0.0002355682518100366
step: 100, loss: 0.0055463663302361965
step: 110, loss: 0.000800244917627424
step: 120, loss: 0.0002972127986140549
step: 130, loss: 0.023939916864037514
step: 140, loss: 0.001537152798846364
step: 150, loss: 0.00731431832537055
step: 160, loss: 0.0002137282572221011
step: 170, loss: 0.0004210870247334242
/usr/local/lib/python3.7/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=0.7461139896373058, f1=0.7735368956743003,
best_f1=0.7957559681697612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012142891995608807
step: 10, loss: 0.03993809223175049
step: 20, loss: 0.000637478893622756
step: 30, loss: 0.00011940652620978653
step: 40, loss: 0.000847768853418529
step: 50, loss: 0.00029426085529848933
step: 60, loss: 0.017026495188474655
step: 70, loss: 0.00027059103013016284
step: 80, loss: 0.0008053270285017788
step: 90, loss: 0.00015387112216558307
step: 100, loss: 0.00038528168806806207
step: 110, loss: 0.00027672480791807175

```

```

step: 120, loss: 0.002644680207595229
step: 130, loss: 0.00014539090625476092
step: 140, loss: 0.0006854782113805413
step: 150, loss: 0.0025327354669570923
step: 160, loss: 0.0006412667571566999
step: 170, loss: 0.006871034391224384
/usr/local/lib/python3.7/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=0.7608695652173912, f1=0.7913279132791328,
best_f1=0.7957559681697612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017939240206032991
step: 10, loss: 0.00012724737462121993
step: 20, loss: 0.00015185891243163496
step: 30, loss: 0.00013427012891042978
step: 40, loss: 0.00017502981063444167
step: 50, loss: 0.00014134906814433634
step: 60, loss: 0.00012228888226673007
step: 70, loss: 0.00015700209769420326
step: 80, loss: 0.0003339051327202469
step: 90, loss: 0.0027553082909435034
step: 100, loss: 0.0033016130328178406
step: 110, loss: 0.00010384165943833068
step: 120, loss: 0.06772724539041519
step: 130, loss: 0.0006785726291127503
step: 140, loss: 0.0006280656671151519
step: 150, loss: 0.00011919863027287647
step: 160, loss: 0.0021097261924296618
step: 170, loss: 0.00037717001396231353
/usr/local/lib/python3.7/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=0.77088948787062, f1=0.8010752688172044,
best_f1=0.7957559681697612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025426523643545806
step: 10, loss: 0.0004889380652457476
step: 20, loss: 0.004963186103850603
step: 30, loss: 0.00014204639592207968
step: 40, loss: 0.0015646623214706779
step: 50, loss: 0.00016685058653820306
step: 60, loss: 0.025563469156622887
step: 70, loss: 0.00014563740114681423

```

```

step: 80, loss: 0.005889447871595621
step: 90, loss: 0.00014653084508609027
step: 100, loss: 0.00023762833734508604
step: 110, loss: 8.133308438118547e-05
step: 120, loss: 0.00011262425687164068
step: 130, loss: 0.0010017314925789833
step: 140, loss: 0.00022569249267689884
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.0003675313200801611
step: 160, loss: 0.00010112916788784787
step: 170, loss: 0.00047753192484378815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7647058823529412, f1=0.806366047745358,
best_f1=0.7957559681697612
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, 245.21it/s]
load_f1 = 0.7759562841530055

```

```

real_f1 = 0.7814207650273223
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, 233.10it/s]

```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

```

[8]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task 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, 566kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 4.10MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 63.8MB/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: 0.6149151921272278
step: 10, loss: 0.6157604455947876
step: 20, loss: 0.49896278977394104
step: 30, loss: 0.32860293984413147
step: 40, loss: 0.21530596911907196
step: 50, loss: 0.029073212295770645
step: 60, loss: 0.0796390026807785
step: 70, loss: 0.03719450533390045
step: 80, loss: 0.049051523208618164
step: 90, loss: 0.16498950123786926
step: 100, loss: 0.009620314463973045
step: 110, loss: 0.3332911431789398
step: 120, loss: 0.012820156291127205
step: 130, loss: 0.10954955220222473
step: 140, loss: 0.03204610198736191
step: 150, loss: 0.07400670647621155
step: 160, loss: 0.03684694692492485

```

```

step: 170, loss: 0.07838038355112076
step: 180, loss: 0.034760475158691406
step: 190, loss: 0.05884244665503502
step: 200, loss: 0.12855884432792664
step: 210, loss: 0.01315909530967474
step: 220, loss: 0.02909490279853344
step: 230, loss: 0.01693233661353588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9617977528089887, f1=0.9641255605381166,
best_f1=0.9641255605381166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022869789972901344
step: 10, loss: 0.10462938249111176
step: 20, loss: 0.19796209037303925
step: 30, loss: 0.1694573312997818
step: 40, loss: 0.07310540974140167
step: 50, loss: 0.014539841562509537
step: 60, loss: 0.0042620557360351086
step: 70, loss: 0.11017307639122009
step: 80, loss: 0.0963747426867485
step: 90, loss: 0.062332406640052795
step: 100, loss: 0.04180927574634552
step: 110, loss: 0.08752545714378357
step: 120, loss: 0.09090100228786469
step: 130, loss: 0.006869492586702108
step: 140, loss: 0.0029354507569223642
step: 150, loss: 0.011863415129482746
step: 160, loss: 0.016987983137369156
step: 170, loss: 0.0006061301683075726
step: 180, loss: 0.13848721981048584
step: 190, loss: 0.055257368832826614
step: 200, loss: 0.028225481510162354
step: 210, loss: 0.003544504288583994
step: 220, loss: 0.009566976688802242
step: 230, loss: 0.04798177629709244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9621380846325166, f1=0.9622222222222223,
best_f1=0.9622222222222223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09455239772796631
step: 10, loss: 0.01825740374624729
step: 20, loss: 0.03781408444046974
step: 30, loss: 0.051320385187864304
step: 40, loss: 0.11732278764247894
step: 50, loss: 0.017408037558197975
step: 60, loss: 0.015465062111616135
step: 70, loss: 0.025335563346743584
step: 80, loss: 0.0023153366055339575
step: 90, loss: 0.11514794081449509
step: 100, loss: 0.0007392295519821346
step: 110, loss: 0.0015571234980598092
step: 120, loss: 0.002156763104721904
step: 130, loss: 0.0005039318348281085
step: 140, loss: 0.00663351034745574
step: 150, loss: 0.009415450505912304
step: 160, loss: 0.23461447656154633
step: 170, loss: 0.11469081044197083
step: 180, loss: 0.041946396231651306
step: 190, loss: 0.014984416775405407
step: 200, loss: 0.014497308060526848
step: 210, loss: 0.05925232172012329
step: 220, loss: 0.0278960932046175
step: 230, loss: 0.0020892454776912928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9658314350797267, f1=0.9681093394077448,
best_f1=0.9681093394077448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005631730891764164
step: 10, loss: 0.001217568526044488
step: 20, loss: 0.0025111373979598284
step: 30, loss: 0.0008844159310683608
step: 40, loss: 0.0015325963031500578

```



```

step: 50, loss: 0.0002706674567889422
step: 60, loss: 0.008224003948271275
step: 70, loss: 0.0011344236554577947
step: 80, loss: 0.0008539056871086359
step: 90, loss: 0.010451119393110275
step: 100, loss: 0.0007176707149483263
step: 110, loss: 0.00046631533768959343
step: 120, loss: 0.007465501315891743
step: 130, loss: 0.007206170354038477
step: 140, loss: 0.0015816095983609557
step: 150, loss: 0.1518407016992569
step: 160, loss: 0.038453128188848495
step: 170, loss: 0.014981027692556381
step: 180, loss: 0.0010065758833661675
step: 190, loss: 0.008912356570363045
step: 200, loss: 0.001578883035108447
step: 210, loss: 0.13200876116752625
step: 220, loss: 0.0005104629672132432
step: 230, loss: 0.07833696901798248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.971815107102593, 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: 0.0005110279307700694
step: 10, loss: 0.0004100445657968521
step: 20, loss: 0.05782388150691986
step: 30, loss: 0.05479517579078674
step: 40, loss: 0.025241967290639877
step: 50, loss: 0.07644841820001602
step: 60, loss: 0.08582468330860138
step: 70, loss: 0.010826285928487778
step: 80, loss: 0.0024159515742212534
step: 90, loss: 0.0019364876206964254
step: 100, loss: 0.0013604938285425305
step: 110, loss: 0.0012465707259252667
step: 120, loss: 0.00865617860108614
step: 130, loss: 0.0010975106852129102
step: 140, loss: 0.007250786758959293
step: 150, loss: 0.0006920408341102302
step: 160, loss: 0.0025436733849346638

```

```

step: 170, loss: 0.055687155574560165
step: 180, loss: 0.0008323714137077332
step: 190, loss: 0.0412159264087677
step: 200, loss: 0.0011693417327478528
step: 210, loss: 0.0005218542646616697
step: 220, loss: 0.0009170316625386477
step: 230, loss: 0.0502018928527832
/usr/local/lib/python3.7/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=0.9694224235560589, f1=0.9797752808988766,
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: 0.008890473283827305
step: 10, loss: 0.002575952559709549
step: 20, loss: 0.0030976193957030773
step: 30, loss: 0.00022896476730238646
step: 40, loss: 0.000265709386439994
step: 50, loss: 0.00042010904871858656
step: 60, loss: 0.0008432003669440746
step: 70, loss: 0.00034848126233555377
step: 80, loss: 0.01746382936835289
step: 90, loss: 0.0008231747196987271
step: 100, loss: 0.03359656780958176
step: 110, loss: 0.03777102753520012
step: 120, loss: 0.001957149710506201
step: 130, loss: 0.00029827351681888103
step: 140, loss: 0.001356559805572033
step: 150, loss: 0.0027922845911234617
step: 160, loss: 0.0017945842118933797
step: 170, loss: 0.0006623430526815355
step: 180, loss: 0.021669279783964157
step: 190, loss: 0.001246413798071444
step: 200, loss: 0.00024402826966252178
step: 210, loss: 0.00452799117192626
step: 220, loss: 0.00031304333242587745
step: 230, loss: 0.05988805741071701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9730941704035874, f1=0.9842342342342343,
best_f1=0.9842342342342343

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015092747053131461
step: 10, loss: 0.00016251049237325788
step: 20, loss: 0.005065700504928827
step: 30, loss: 0.0001669956836849451
step: 40, loss: 0.00025292523787356913
step: 50, loss: 0.0004377156146802008
step: 60, loss: 0.00017552934878040105
step: 70, loss: 0.08622075617313385
step: 80, loss: 0.007468738127499819
step: 90, loss: 8.738309406908229e-05
step: 100, loss: 0.0002339666389161721
step: 110, loss: 0.00014324243238661438
step: 120, loss: 0.00011252309195697308
step: 130, loss: 0.00017427746206521988
step: 140, loss: 0.00036476648529060185
step: 150, loss: 0.0004878986510448158
step: 160, loss: 0.025045771151781082
step: 170, loss: 0.0027980993036180735
step: 180, loss: 0.0006236847839318216
step: 190, loss: 0.00038604048313573003
step: 200, loss: 0.029536070302128792
step: 210, loss: 0.0001246351603185758
step: 220, loss: 0.000990306376479566
step: 230, loss: 0.0012232891749590635
/usr/local/lib/python3.7/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=0.968609865470852, f1=0.9753363228699552,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005807020934298635
step: 10, loss: 0.0037615762557834387
step: 20, loss: 0.00018793570052366704
step: 30, loss: 0.000210518017411232
step: 40, loss: 0.0028293782379478216
step: 50, loss: 0.0036455111112445593
step: 60, loss: 0.0002691514091566205
step: 70, loss: 0.0002476766530890018
step: 80, loss: 0.0011632053647190332
step: 90, loss: 0.00010551418381510302
step: 100, loss: 0.000576366379391402
step: 110, loss: 0.12386876344680786
step: 120, loss: 0.0026103754062205553

```

```

step: 130, loss: 0.0011880015954375267
step: 140, loss: 0.00018622119387146086
step: 150, loss: 0.00017118534015025944
step: 160, loss: 0.00043792877113446593
step: 170, loss: 0.036523688584566116
step: 180, loss: 0.008909414522349834
step: 190, loss: 0.06747126579284668
step: 200, loss: 0.00016515323659405112
step: 210, loss: 0.0017101522535085678
step: 220, loss: 0.0004229988844599575
step: 230, loss: 0.0011381716467440128
/usr/local/lib/python3.7/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=0.9697648376259798, f1=0.972129319955407,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018154091958422214
step: 10, loss: 0.0027412448544055223
step: 20, loss: 0.00047335962881334126
step: 30, loss: 6.789527833461761e-05
step: 40, loss: 0.013827181421220303
step: 50, loss: 0.00038439975469373167
step: 60, loss: 0.0020947833545506
step: 70, loss: 9.65762956184335e-05
step: 80, loss: 0.00014533383364323527
step: 90, loss: 0.0004019986081402749
step: 100, loss: 0.00018692782032303512
step: 110, loss: 6.687956192763522e-05
step: 120, loss: 5.608419087366201e-05
step: 130, loss: 0.00018417200772091746
step: 140, loss: 8.082435670075938e-05
step: 150, loss: 0.0013088196283206344
step: 160, loss: 5.339454946806654e-05
step: 170, loss: 0.00018721104424912483
step: 180, loss: 0.000766140699852258
step: 190, loss: 0.0002704434155020863
step: 200, loss: 0.0001174425779026933
step: 210, loss: 0.00013837922597303987
step: 220, loss: 4.767367136082612e-05
step: 230, loss: 0.00037635283661074936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9741282339707535, f1=0.9752252252252253,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002638149890117347
step: 10, loss: 0.00016553536988794804
step: 20, loss: 4.868226096732542e-05
step: 30, loss: 9.274829062633216e-05
step: 40, loss: 7.356169953709468e-05
step: 50, loss: 5.6704491726122797e-05
step: 60, loss: 0.00017264235066249967
step: 70, loss: 0.0002908711030613631
step: 80, loss: 8.8931767095346e-05
step: 90, loss: 0.0002569802454672754
step: 100, loss: 4.6418284910032526e-05
step: 110, loss: 7.845889194868505e-05
step: 120, loss: 0.00028941742493771017
step: 130, loss: 0.00010317245323676616
step: 140, loss: 0.042606353759765625
step: 150, loss: 0.036430664360523224
step: 160, loss: 7.491608994314447e-05
step: 170, loss: 4.853700374951586e-05
step: 180, loss: 0.0013542554806917906
step: 190, loss: 0.002299710176885128
step: 200, loss: 3.944872514693998e-05
step: 210, loss: 0.00031708364258520305
step: 220, loss: 8.223443728638813e-05
step: 230, loss: 0.0002786130062304437
/usr/local/lib/python3.7/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=0.9707207207207207, f1=0.9763779527559054,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.445512942969799e-05
step: 10, loss: 0.00013229349860921502
step: 20, loss: 0.0001612775376997888
step: 30, loss: 0.0006086168577894568
step: 40, loss: 0.00011980049748672172
step: 50, loss: 5.608334322459996e-05
step: 60, loss: 0.011346694082021713
step: 70, loss: 0.00031005532946437597
step: 80, loss: 3.5783712519332767e-05

```

```

step: 90, loss: 8.930317562771961e-05
step: 100, loss: 8.217254071496427e-05
step: 110, loss: 0.0013939541531726718
step: 120, loss: 4.9885718908626586e-05
step: 130, loss: 6.067248614272103e-05
step: 140, loss: 4.8204849008470774e-05
step: 150, loss: 0.015378871001303196
step: 160, loss: 2.9238979550427757e-05
step: 170, loss: 0.00043452956015244126
step: 180, loss: 0.0001157947481260635
step: 190, loss: 3.423012822167948e-05
step: 200, loss: 0.00011272386473137885
step: 210, loss: 3.720622407854535e-05
step: 220, loss: 4.197440648567863e-05
step: 230, loss: 4.969084693584591e-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=0.970917225950783, f1=0.9753914988814317,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.0433743783505633e-05
step: 10, loss: 4.673085277318023e-05
step: 20, loss: 6.395357922883704e-05
step: 30, loss: 8.439907833235338e-05
step: 40, loss: 0.00040497854934073985
step: 50, loss: 0.0002615534176584333
step: 60, loss: 0.00013161909009795636
step: 70, loss: 5.8803077990887687e-05
step: 80, loss: 0.00011117149551864713
step: 90, loss: 3.0233513825805858e-05
step: 100, loss: 3.214041134924628e-05
step: 110, loss: 3.7806144973728806e-05
step: 120, loss: 4.0352493670070544e-05
step: 130, loss: 3.6900695704389364e-05
step: 140, loss: 0.0007155967177823186
step: 150, loss: 4.249274206813425e-05
step: 160, loss: 5.0924543756991625e-05
step: 170, loss: 3.225240288884379e-05
step: 180, loss: 0.005162614397704601
step: 190, loss: 0.00012887213961221278
step: 200, loss: 3.09184288198594e-05
step: 210, loss: 5.4305040976032615e-05
step: 220, loss: 4.0456892747897655e-05
step: 230, loss: 0.025572674348950386
/usr/local/lib/python3.7/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=0.9730941704035874, f1=0.9731543624161074,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.704731691163033e-05
step: 10, loss: 0.00011407359124859795
step: 20, loss: 9.23673651413992e-05
step: 30, loss: 0.00010561208910075948
step: 40, loss: 8.4567756857723e-05
step: 50, loss: 6.894624675624073e-05
step: 60, loss: 3.916191781172529e-05
step: 70, loss: 3.125378498225473e-05
step: 80, loss: 3.5213553928770125e-05
step: 90, loss: 6.938260048627853e-05
step: 100, loss: 3.917621143045835e-05
step: 110, loss: 0.005998645443469286
step: 120, loss: 0.014253332279622555
step: 130, loss: 3.798888064920902e-05
step: 140, loss: 5.951978891971521e-05
step: 150, loss: 6.960642349440604e-05
step: 160, loss: 0.00022091748542152345
step: 170, loss: 7.030621054582298e-05
step: 180, loss: 4.679746416513808e-05
step: 190, loss: 3.4940807381644845e-05
step: 200, loss: 5.144074748386629e-05
step: 210, loss: 2.981217585329432e-05
step: 220, loss: 3.809266854659654e-05
step: 230, loss: 3.312391345389187e-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=0.9705882352941176, f1=0.9784824462061155,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.0886617827927694e-05
step: 10, loss: 7.948796701384708e-05
step: 20, loss: 4.48657083325088e-05
step: 30, loss: 4.235079904901795e-05
step: 40, loss: 0.021142516285181046
step: 50, loss: 4.624296343536116e-05
step: 60, loss: 4.365641507320106e-05
step: 70, loss: 6.462140299845487e-05
step: 80, loss: 4.236547101754695e-05

```

```

step: 90, loss: 6.598773325094953e-05
step: 100, loss: 4.548646757029928e-05
step: 110, loss: 0.0002736106689553708
step: 120, loss: 2.8467278752941638e-05
step: 130, loss: 4.8462195991305634e-05
step: 140, loss: 3.9672300772508606e-05
step: 150, loss: 5.047932063462213e-05
step: 160, loss: 0.003619818016886711
step: 170, loss: 1.714728568913415e-05
step: 180, loss: 0.00010102267697220668
step: 190, loss: 0.00010253037908114493
step: 200, loss: 3.634511449490674e-05
step: 210, loss: 4.318798892199993e-05
step: 220, loss: 0.0004602159315254539
step: 230, loss: 0.0004352243850007653
/usr/local/lib/python3.7/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=0.9706546275395034, f1=0.9796380090497738,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.617669088067487e-05
step: 10, loss: 2.1010113414376974e-05
step: 20, loss: 3.998868851340376e-05
step: 30, loss: 5.4933741921558976e-05
step: 40, loss: 8.861716196406633e-05
step: 50, loss: 0.014391402713954449
step: 60, loss: 5.862842226633802e-05
step: 70, loss: 6.43418388790451e-05
step: 80, loss: 4.037117832922377e-05
step: 90, loss: 8.286676893476397e-05
step: 100, loss: 3.3526313927723095e-05
step: 110, loss: 4.7220277338055894e-05
step: 120, loss: 3.583939542295411e-05
step: 130, loss: 0.002645030850544572
step: 140, loss: 4.5328524720389396e-05
step: 150, loss: 5.911013431614265e-05
step: 160, loss: 2.22170892811846e-05
step: 170, loss: 3.973702769144438e-05
step: 180, loss: 0.00021563551854342222
step: 190, loss: 3.581684723030776e-05
step: 200, loss: 2.8758440748788416e-05
step: 210, loss: 4.55156150565017e-05
step: 220, loss: 0.00032544598798267543
step: 230, loss: 3.12539232254494e-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=0.9706546275395034, f1=0.9785310734463276,
best_f1=0.9752252252252253
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
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, 194.02it/s]
load_f1 = 0.9740698985343857
real_f1 = 0.9741282339707535
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, 231.28it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6496710181236267
step: 10, loss: 0.5355506539344788
step: 20, loss: 0.5474937558174133
step: 30, loss: 0.1977870911359787
step: 40, loss: 0.199203759431839
step: 50, loss: 0.22630909085273743
step: 60, loss: 0.11098060756921768
step: 70, loss: 0.11724264919757843
step: 80, loss: 0.06727444380521774
step: 90, loss: 0.21277949213981628
step: 100, loss: 0.06448571383953094
step: 110, loss: 0.09347495436668396
step: 120, loss: 0.09402161091566086
step: 130, loss: 0.05656568333506584
step: 140, loss: 0.06281556934118271
step: 150, loss: 0.029476486146450043
step: 160, loss: 0.03431320935487747
step: 170, loss: 0.2646041214466095
step: 180, loss: 0.058851458132267
step: 190, loss: 0.02995334006845951
step: 200, loss: 0.12394876778125763
step: 210, loss: 0.118414968252182
step: 220, loss: 0.22789721190929413
step: 230, loss: 0.16635273396968842
step: 240, loss: 0.14570163190364838
step: 250, loss: 0.03691399097442627
step: 260, loss: 0.05003497377038002
step: 270, loss: 0.022751180455088615
step: 280, loss: 0.10740266740322113

```

```

step: 290, loss: 0.07797804474830627
step: 300, loss: 0.047100815922021866
step: 310, loss: 0.31003692746162415
step: 320, loss: 0.151001438498497
step: 330, loss: 0.0743824690580368
step: 340, loss: 0.07431641221046448
step: 350, loss: 0.07062654942274094
step: 360, loss: 0.18231913447380066
step: 370, loss: 0.07309827208518982
step: 380, loss: 0.02723441645503044
step: 390, loss: 0.143576517701149
step: 400, loss: 0.2597905099391937
step: 410, loss: 0.09891984611749649
step: 420, loss: 0.030317245051264763
step: 430, loss: 0.10222964733839035
step: 440, loss: 0.020777028053998947
step: 450, loss: 0.03128473088145256
step: 460, loss: 0.016567949205636978
step: 470, loss: 0.10734525322914124
step: 480, loss: 0.12771886587142944
step: 490, loss: 0.07603038102388382
step: 500, loss: 0.1711377501487732
step: 510, loss: 0.05575227737426758
step: 520, loss: 0.13102181255817413
step: 530, loss: 0.004091911017894745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9246919214970334, f1=0.9249771271729187,
best_f1=0.9249771271729187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12402001768350601
step: 10, loss: 0.08282976597547531
step: 20, loss: 0.011573459021747112
step: 30, loss: 0.02298850193619728
step: 40, loss: 0.04273645579814911
step: 50, loss: 0.19738708436489105
step: 60, loss: 0.011571216396987438
step: 70, loss: 0.03472288325428963
step: 80, loss: 0.07181359827518463
step: 90, loss: 0.014796201139688492
step: 100, loss: 0.019887851551175117

```

```
step: 110, loss: 0.026556219905614853
step: 120, loss: 0.017837906256318092
step: 130, loss: 0.07103241235017776
step: 140, loss: 0.08637451380491257
step: 150, loss: 0.02713618613779545
step: 160, loss: 0.04622432217001915
step: 170, loss: 0.12571386992931366
step: 180, loss: 0.050673194229602814
step: 190, loss: 0.07489009946584702
step: 200, loss: 0.007570087444037199
step: 210, loss: 0.03942751884460449
step: 220, loss: 0.025336802005767822
step: 230, loss: 0.030194880440831184
step: 240, loss: 0.02168498933315277
step: 250, loss: 0.08568920195102692
step: 260, loss: 0.03307507932186127
step: 270, loss: 0.1979183852672577
step: 280, loss: 0.11827126145362854
step: 290, loss: 0.029966123402118683
step: 300, loss: 0.19248470664024353
step: 310, loss: 0.009243116714060307
step: 320, loss: 0.1541922241449356
step: 330, loss: 0.06171722337603569
step: 340, loss: 0.015498477034270763
step: 350, loss: 0.006829812191426754
step: 360, loss: 0.1282108575105667
step: 370, loss: 0.10787132382392883
step: 380, loss: 0.06798841059207916
step: 390, loss: 0.0438622310757637
step: 400, loss: 0.1245722770690918
step: 410, loss: 0.04703742265701294
step: 420, loss: 0.11779433488845825
step: 430, loss: 0.053284477442502975
step: 440, loss: 0.17642110586166382
step: 450, loss: 0.13128606975078583
step: 460, loss: 0.04054923728108406
step: 470, loss: 0.08555746078491211
step: 480, loss: 0.19525031745433807
step: 490, loss: 0.02593224123120308
step: 500, loss: 0.18956823647022247
step: 510, loss: 0.011702745221555233
step: 520, loss: 0.0774252712726593
step: 530, loss: 0.01239960640668869
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9260450160771704, f1=0.9284403669724771,
best_f1=0.9284403669724771
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12761543691158295
step: 10, loss: 0.02692047506570816
step: 20, loss: 0.04710455238819122
step: 30, loss: 0.07212059944868088
step: 40, loss: 0.19531728327274323
step: 50, loss: 0.12339923530817032
step: 60, loss: 0.023849422112107277
step: 70, loss: 0.0027424136642366648
step: 80, loss: 0.017781171947717667
step: 90, loss: 0.006694050505757332
step: 100, loss: 0.12935857474803925
step: 110, loss: 0.0039881146512925625
step: 120, loss: 0.004681040532886982
step: 130, loss: 0.020981481298804283
step: 140, loss: 0.020270463079214096
step: 150, loss: 0.013434026390314102
step: 160, loss: 0.007404050324112177
step: 170, loss: 0.014107425697147846
step: 180, loss: 0.009945429861545563
step: 190, loss: 0.012379590421915054
step: 200, loss: 0.04726565629243851
step: 210, loss: 0.09529230743646622
step: 220, loss: 0.024814285337924957
step: 230, loss: 0.05600526183843613
step: 240, loss: 0.014311379753053188
step: 250, loss: 0.03650589659810066
step: 260, loss: 0.022331558167934418
step: 270, loss: 0.008300729095935822
step: 280, loss: 0.060539379715919495
step: 290, loss: 0.0026584502775222063
step: 300, loss: 0.06238330155611038
step: 310, loss: 0.008183641359210014
step: 320, loss: 0.022303259000182152
step: 330, loss: 0.009579619392752647
step: 340, loss: 0.002222067443653941
step: 350, loss: 0.0033683988731354475
step: 360, loss: 0.05438629910349846
step: 370, loss: 0.016209183260798454
step: 380, loss: 0.015004412271082401
step: 390, loss: 0.017906207591295242
step: 400, loss: 0.00947515107691288

```

```

step: 410, loss: 0.005249433685094118
step: 420, loss: 0.10279760509729385
step: 430, loss: 0.015238862484693527
step: 440, loss: 0.009358610026538372
step: 450, loss: 0.05511041358113289
step: 460, loss: 0.033059991896152496
step: 470, loss: 0.07353335618972778
step: 480, loss: 0.004735864233225584
step: 490, loss: 0.005474486853927374
step: 500, loss: 0.019051043316721916
step: 510, loss: 0.02098867855966091
step: 520, loss: 0.05631197616457939
step: 530, loss: 0.024868082255125046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9310344827586207, f1=0.9225023342670402,
best_f1=0.9225023342670402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01980866864323616
step: 10, loss: 0.005490655545145273
step: 20, loss: 0.005011112429201603
step: 30, loss: 0.0011700086761265993
step: 40, loss: 0.0028695114888250828
step: 50, loss: 0.017948588356375694
step: 60, loss: 0.0022288148757070303
step: 70, loss: 0.009024304337799549
step: 80, loss: 0.014230922795832157
step: 90, loss: 0.07210805267095566
step: 100, loss: 0.0016932712169364095
step: 110, loss: 0.005408769007772207
step: 120, loss: 0.003618905320763588
step: 130, loss: 0.0008444340783171356
step: 140, loss: 0.000744595134165138
step: 150, loss: 0.0014196683187037706
step: 160, loss: 0.07691176235675812
step: 170, loss: 0.005852863658219576
step: 180, loss: 0.0046478332951664925
step: 190, loss: 0.08499796688556671
step: 200, loss: 0.007401813752949238
step: 210, loss: 0.044213343411684036
step: 220, loss: 0.0064554098062217236

```

```

step: 230, loss: 0.029861368238925934
step: 240, loss: 0.004553196486085653
step: 250, loss: 0.033087991178035736
step: 260, loss: 0.009218159131705761
step: 270, loss: 0.006069767288863659
step: 280, loss: 0.11987975984811783
step: 290, loss: 0.034156739711761475
step: 300, loss: 0.0009717888315208256
step: 310, loss: 0.09984275698661804
step: 320, loss: 0.0021901768632233143
step: 330, loss: 0.006877771578729153
step: 340, loss: 0.1298915594816208
step: 350, loss: 0.006486755795776844
step: 360, loss: 0.008373936638236046
step: 370, loss: 0.004640497267246246
step: 380, loss: 0.0028901954647153616
step: 390, loss: 0.003491985844448209
step: 400, loss: 0.0007395677966997027
step: 410, loss: 0.0028256536461412907
step: 420, loss: 0.05533720925450325
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 0.2986679673194885
step: 440, loss: 0.047281503677368164
step: 450, loss: 0.03140322491526604
step: 460, loss: 0.02597256377339363
step: 470, loss: 0.013161000795662403
step: 480, loss: 0.041553109884262085
step: 490, loss: 0.011153471656143665
step: 500, loss: 0.0029792541172355413
step: 510, loss: 0.01649804785847664
step: 520, loss: 0.036048270761966705
step: 530, loss: 0.05213514715433121
/usr/local/lib/python3.7/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=0.9236430542778289, f1=0.9301470588235294,
best_f1=0.9225023342670402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03914127126336098
step: 10, loss: 0.05313700810074806
step: 20, loss: 0.003619518829509616
step: 30, loss: 0.0025986114051193
step: 40, loss: 0.001701953005976975
step: 50, loss: 0.0849851742386818
step: 60, loss: 0.07474177330732346
step: 70, loss: 0.004792123567312956

```


step: 80, loss: 0.002728985855355859
step: 90, loss: 0.08227050304412842
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.003499666228890419
step: 110, loss: 0.0013891300186514854
step: 120, loss: 0.0014662927715107799
step: 130, loss: 0.0003974221181124449
step: 140, loss: 0.0014383683446794748
step: 150, loss: 0.0028773059602826834
step: 160, loss: 0.00223339325748384
step: 170, loss: 0.004463139455765486
step: 180, loss: 0.008507385849952698
step: 190, loss: 0.0008084392175078392
step: 200, loss: 0.0008547122124582529
step: 210, loss: 0.010911817662417889
step: 220, loss: 0.013245632871985435
step: 230, loss: 0.0059165144339203835
step: 240, loss: 0.004092320799827576
step: 250, loss: 0.02236495353281498
step: 260, loss: 0.027840912342071533
step: 270, loss: 0.005176469683647156
step: 280, loss: 0.019035179167985916
step: 290, loss: 0.1346452385187149
step: 300, loss: 0.0008709232788532972
step: 310, loss: 0.015449218451976776
step: 320, loss: 0.08903078734874725
step: 330, loss: 0.01867162249982357
step: 340, loss: 0.0026656093541532755
step: 350, loss: 0.0023844523821026087
step: 360, loss: 0.01517312042415142
step: 370, loss: 0.02403443120419979
step: 380, loss: 0.000515331921633333
step: 390, loss: 0.00020528071036096662
step: 400, loss: 0.002189124235883355
step: 410, loss: 0.002498648827895522
step: 420, loss: 0.005367943085730076
step: 430, loss: 0.003955364692956209
step: 440, loss: 0.010672789067029953
step: 450, loss: 0.019898006692528725
step: 460, loss: 0.007720068097114563
step: 470, loss: 0.0024507285561412573
step: 480, loss: 0.005106728989630938
step: 490, loss: 0.00015589609392918646
step: 500, loss: 0.0004854470316786319
step: 510, loss: 0.22279532253742218
step: 520, loss: 0.011761325411498547
step: 530, loss: 0.003931534010916948
/usr/local/lib/python3.7/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=0.9261495587552253, f1=0.9212855146716348,
best_f1=0.9225023342670402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008508889586664736
step: 10, loss: 0.00017819119966588914
step: 20, loss: 0.02238820306956768
step: 30, loss: 0.0005004773265682161
step: 40, loss: 0.002278541214764118
step: 50, loss: 0.2002110332250595
step: 60, loss: 8.722676284378394e-05
step: 70, loss: 0.0003555336152203381
step: 80, loss: 0.01965332217514515
step: 90, loss: 0.00018340650422032923
step: 100, loss: 0.004271746147423983
step: 110, loss: 0.03439374640583992
step: 120, loss: 0.022239815443754196
step: 130, loss: 0.00035815531737171113
step: 140, loss: 0.0008803320815786719
step: 150, loss: 0.0013662419514730573
step: 160, loss: 0.0024828920140862465
step: 170, loss: 0.0003286396968178451
step: 180, loss: 0.0003070829261559993
step: 190, loss: 0.0026877657510340214
step: 200, loss: 0.0024812619667500257
step: 210, loss: 0.0005910171894356608
step: 220, loss: 0.0026797230821102858
step: 230, loss: 0.0002254528080811724
step: 240, loss: 0.011726761236786842
step: 250, loss: 0.004729270935058594
step: 260, loss: 9.404408774571493e-05
step: 270, loss: 0.11738213151693344
step: 280, loss: 0.0025060181505978107
step: 290, loss: 0.0010880957124754786
step: 300, loss: 0.000519168097525835
step: 310, loss: 0.00029881164664402604
step: 320, loss: 0.0002959598496090621
step: 330, loss: 0.00013450307596940547
step: 340, loss: 0.03729816898703575
step: 350, loss: 0.0026700361631810665
step: 360, loss: 0.027725275605916977
step: 370, loss: 0.0008200588054023683
step: 380, loss: 7.75385633460246e-05
step: 390, loss: 0.1773785948753357
step: 400, loss: 0.049010019749403

```

```

step: 410, loss: 0.0003553417627699673
step: 420, loss: 0.0028727829921990633
step: 430, loss: 0.0005701988702639937
step: 440, loss: 0.00040749588515609503
step: 450, loss: 0.0017815523315221071
step: 460, loss: 0.00018915708642452955
step: 470, loss: 0.001427693059667945
step: 480, loss: 0.0005052773631177843
step: 490, loss: 0.005355267319828272
step: 500, loss: 0.004725165199488401
step: 510, loss: 0.0037835370749235153
step: 520, loss: 0.015028182417154312
step: 530, loss: 0.005576055962592363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9325789721829325, f1=0.9163533834586467,
best_f1=0.9163533834586467
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014017593348398805
step: 10, loss: 0.0009875638643279672
step: 20, loss: 0.00015379091200884432
step: 30, loss: 0.013843104243278503
step: 40, loss: 6.430042412830517e-05
step: 50, loss: 0.001847312436439097
step: 60, loss: 0.0013850578106939793
step: 70, loss: 6.979506724746898e-05
step: 80, loss: 0.0012918765423819423
step: 90, loss: 9.303769184043631e-05
step: 100, loss: 0.004116305150091648
step: 110, loss: 0.02110953815281391
step: 120, loss: 0.002181330928578973
step: 130, loss: 0.0008316265884786844
step: 140, loss: 4.90560159960296e-05
step: 150, loss: 0.00013529883290175349
step: 160, loss: 0.00013500881323125213
step: 170, loss: 0.00013305128959473222
step: 180, loss: 0.00011801895016105846
step: 190, loss: 0.0006034215330146253
step: 200, loss: 0.016331614926457405
step: 210, loss: 0.006287342868745327
step: 220, loss: 9.591260459274054e-05

```

```

step: 230, loss: 0.0007051219581626356
step: 240, loss: 0.00023784601944498718
step: 250, loss: 0.004091073293238878
step: 260, loss: 0.0064830174669623375
step: 270, loss: 0.0015235681785270572
step: 280, loss: 0.0009280066005885601
step: 290, loss: 0.00012387344031594694
step: 300, loss: 0.02644573152065277
step: 310, loss: 0.00040870060911402106
step: 320, loss: 0.0016001766780391335
step: 330, loss: 0.0011514101643115282
step: 340, loss: 0.021427685394883156
step: 350, loss: 0.06835515797138214
step: 360, loss: 0.008316755294799805
step: 370, loss: 0.0013797470601275563
step: 380, loss: 0.0007066255784593523
step: 390, loss: 0.00020489796588663012
step: 400, loss: 0.00043161294888705015
step: 410, loss: 0.0027975074481219053
step: 420, loss: 0.000234353487030603
step: 430, loss: 0.010694285854697227
step: 440, loss: 0.0014771883143112063
step: 450, loss: 0.0006938835140317678
step: 460, loss: 0.0011100288247689605
step: 470, loss: 0.0010099916253238916
step: 480, loss: 0.10705417394638062
step: 490, loss: 0.00033843907294794917
step: 500, loss: 0.001751324743963778
step: 510, loss: 0.0004784624034073204
step: 520, loss: 0.0028669950552284718
step: 530, loss: 0.06612148880958557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9355900329102022, f1=0.9276315789473684,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017169548198580742
step: 10, loss: 0.07758825272321701
step: 20, loss: 0.015490395948290825
step: 30, loss: 0.001036180998198688
step: 40, loss: 0.00010359942825743929

```

step: 50, loss: 0.0078091928735375404
step: 60, loss: 0.0034609397407621145
step: 70, loss: 0.0035546543076634407
step: 80, loss: 0.0003837542317342013
step: 90, loss: 0.0002869471500162035
step: 100, loss: 5.13641134602949e-05
step: 110, loss: 9.581098129274324e-05
step: 120, loss: 7.942014781292528e-05
step: 130, loss: 8.752122084842995e-05
step: 140, loss: 0.0002638769510667771
step: 150, loss: 7.355814886977896e-05
step: 160, loss: 0.00027201493503525853
step: 170, loss: 0.0031046187505126
step: 180, loss: 3.543757338775322e-05
step: 190, loss: 0.0007692381623201072
step: 200, loss: 0.0011586080072447658
step: 210, loss: 0.0003060781746171415
step: 220, loss: 0.0013601260725408792
step: 230, loss: 4.3335752707207575e-05
step: 240, loss: 4.004157744930126e-05
step: 250, loss: 0.0010718742851167917
step: 260, loss: 0.00015127180085983127
step: 270, loss: 0.0003971937985625118
step: 280, loss: 0.018752234056591988
step: 290, loss: 0.0008885873248800635
step: 300, loss: 8.209356019506231e-05
step: 310, loss: 0.0002597779966890812
step: 320, loss: 0.06893672794103622
step: 330, loss: 0.0004869972472079098
step: 340, loss: 0.0002921413688454777
step: 350, loss: 0.0014933559577912092
step: 360, loss: 0.029668506234884262
step: 370, loss: 0.0090717738494277
step: 380, loss: 0.0029295661952346563
step: 390, loss: 0.13116835057735443
step: 400, loss: 0.0009866226464509964
step: 410, loss: 0.00476897181943059
step: 420, loss: 0.0028756738174706697
step: 430, loss: 0.023632198572158813
step: 440, loss: 0.009295489639043808
step: 450, loss: 0.000703193771187216
step: 460, loss: 0.0004679822304751724
step: 470, loss: 0.0003602505603339523
step: 480, loss: 0.00313800061121583
step: 490, loss: 0.0005164105095900595
step: 500, loss: 0.0004383153573144227
step: 510, loss: 0.01997150480747223
step: 520, loss: 0.002779472852125764

```
step: 530, loss: 0.0009291535243391991
/usr/local/lib/python3.7/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=0.9322892676186089, f1=0.9256198347107438,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009005204774439335
step: 10, loss: 9.1419045929797e-05
step: 20, loss: 0.0010289368219673634
step: 30, loss: 0.0005577282281592488
step: 40, loss: 0.0005165900220163167
step: 50, loss: 0.0005334774032235146
step: 60, loss: 8.144017192535102e-05
step: 70, loss: 6.534229760291055e-05
step: 80, loss: 0.0037808988709002733
step: 90, loss: 0.00013302791921887547
step: 100, loss: 4.8399775550933555e-05
step: 110, loss: 0.00014693886623717844
step: 120, loss: 0.0002597273560240865
step: 130, loss: 0.003588299499824643
step: 140, loss: 0.0006178840412758291
step: 150, loss: 9.660929208621383e-05
step: 160, loss: 0.0006614931626245379
step: 170, loss: 0.012596135027706623
step: 180, loss: 0.00034493140992708504
step: 190, loss: 5.82555185246747e-05
step: 200, loss: 0.00011138707486679778
step: 210, loss: 0.00013515220780391246
step: 220, loss: 0.006519306916743517
step: 230, loss: 4.0060069295577705e-05
step: 240, loss: 6.980508624110371e-05
step: 250, loss: 0.0006777657545171678
step: 260, loss: 0.0002510215272195637
step: 270, loss: 3.6533292586682364e-05
step: 280, loss: 7.619534153491259e-05
step: 290, loss: 0.07094468921422958
step: 300, loss: 6.703515828121454e-05
step: 310, loss: 0.002071283059194684
step: 320, loss: 0.0002521497954148799
step: 330, loss: 0.0006217581103555858
step: 340, loss: 0.002944893203675747
step: 350, loss: 0.00016839808085933328
step: 360, loss: 0.0022739616688340902
step: 370, loss: 0.006566685624420643
step: 380, loss: 0.0005842982791364193
```

```

step: 390, loss: 0.00037295836955308914
step: 400, loss: 0.00010332548117730767
step: 410, loss: 2.384881918260362e-05
step: 420, loss: 3.06059155263938e-05
step: 430, loss: 2.8250760806258768e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.037789005786180496
step: 450, loss: 5.444569978863001e-05
step: 460, loss: 0.0019329945789650083
step: 470, loss: 2.9476943382178433e-05
step: 480, loss: 3.093729173997417e-05
step: 490, loss: 0.0012321891263127327
step: 500, loss: 0.00033979411819018424
step: 510, loss: 0.05145140737295151
step: 520, loss: 0.0001224612060468644
step: 530, loss: 4.8664267524145544e-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=0.9344490934449092, f1=0.9229340761374188,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001497392076998949
step: 10, loss: 4.625748988473788e-05
step: 20, loss: 4.461314529180527e-05
step: 30, loss: 0.0012541132746264338
step: 40, loss: 0.0001417471212334931
step: 50, loss: 6.785122968722135e-05
step: 60, loss: 0.0009081550524570048
step: 70, loss: 7.220457337098196e-05
step: 80, loss: 4.829062891076319e-05
step: 90, loss: 0.033630553632974625
step: 100, loss: 0.00012416404206305742
step: 110, loss: 5.946851160842925e-05
step: 120, loss: 0.0003380376729182899
step: 130, loss: 6.114449934102595e-05
step: 140, loss: 0.007971656508743763
step: 150, loss: 7.46168807381764e-05
step: 160, loss: 5.830423106090166e-05
step: 170, loss: 8.107630856102332e-05
step: 180, loss: 0.00014402194938156754
step: 190, loss: 0.00197012210264802
step: 200, loss: 0.020723868161439896
step: 210, loss: 0.00010227707389276475
step: 220, loss: 0.00043638679198920727
step: 230, loss: 0.002891377080231905

```

```

step: 240, loss: 0.0001657611137488857
step: 250, loss: 8.540226554032415e-05
step: 260, loss: 0.0017233610851690173
step: 270, loss: 0.001260785385966301
step: 280, loss: 0.00201358157210052
step: 290, loss: 5.1293009164510295e-05
step: 300, loss: 2.6038695068564266e-05
step: 310, loss: 6.695985211990774e-05
step: 320, loss: 0.0019887685775756836
step: 330, loss: 0.02897321805357933
step: 340, loss: 0.0009510593372397125
step: 350, loss: 0.008703375235199928
step: 360, loss: 0.00021042714070063084
step: 370, loss: 0.0012009673519060016
step: 380, loss: 0.02265334129333496
step: 390, loss: 0.01635158248245716
step: 400, loss: 0.0019270532066002488
step: 410, loss: 0.0013415950816124678
step: 420, loss: 0.0003838356351479888
step: 430, loss: 0.00011974716471740976
step: 440, loss: 0.00045630341628566384
step: 450, loss: 0.00030761692323721945
step: 460, loss: 0.0007688611513003707
step: 470, loss: 0.0002251885598525405
step: 480, loss: 4.405945219332352e-05
step: 490, loss: 0.00013110156578477472
step: 500, loss: 0.00010617743100738153
step: 510, loss: 2.8259346436243504e-05
step: 520, loss: 4.9564889195607975e-05
step: 530, loss: 0.0021473721135407686
/usr/local/lib/python3.7/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=0.9330783938814532, f1=0.9238820171265462,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001189882168546319
step: 10, loss: 0.00017495284555479884
step: 20, loss: 5.342162330634892e-05
step: 30, loss: 0.0001489101559855044
step: 40, loss: 0.0005786457913927734
step: 50, loss: 5.485421934281476e-05
step: 60, loss: 0.00026989573962055147
step: 70, loss: 3.956340151489712e-05
step: 80, loss: 3.509642192511819e-05
step: 90, loss: 4.718109266832471e-05

```



```
step: 100, loss: 0.00045152020175009966
step: 110, loss: 0.0005811608862131834
step: 120, loss: 2.5908298994181678e-05
step: 130, loss: 2.9866492695873603e-05
step: 140, loss: 2.7427242457633838e-05
step: 150, loss: 2.4049119019764476e-05
step: 160, loss: 9.998201858252287e-05
step: 170, loss: 3.3622847695369273e-05
step: 180, loss: 0.0002522044233046472
step: 190, loss: 0.00030286627588793635
step: 200, loss: 4.639502731151879e-05
step: 210, loss: 0.0002724077785387635
step: 220, loss: 0.00010435856529511511
step: 230, loss: 1.7083835700759664e-05
step: 240, loss: 0.00016867340309545398
step: 250, loss: 4.941126826452091e-05
step: 260, loss: 8.130529022309929e-05
step: 270, loss: 0.00012771868205163628
step: 280, loss: 0.00015172007260844111
step: 290, loss: 0.0007631753105670214
step: 300, loss: 4.279134373064153e-05
step: 310, loss: 0.05873999744653702
step: 320, loss: 3.977911183028482e-05
step: 330, loss: 0.0001933604944497347
step: 340, loss: 6.665341788902879e-05
step: 350, loss: 4.796592838829383e-05
step: 360, loss: 0.00011937168892472982
step: 370, loss: 0.0024962674360722303
step: 380, loss: 0.00038275550468824804
step: 390, loss: 0.0013989057624712586
step: 400, loss: 0.001894546439871192
step: 410, loss: 2.59309854300227e-05
step: 420, loss: 0.0007097817142494023
step: 430, loss: 9.290764864999801e-05
step: 440, loss: 6.309049786068499e-05
step: 450, loss: 5.073110878583975e-05
step: 460, loss: 0.0036855945363640785
step: 470, loss: 8.953203359851614e-05
step: 480, loss: 0.0007602816913276911
step: 490, loss: 0.00020592061628121883
step: 500, loss: 8.990702190203592e-05
step: 510, loss: 3.982999623985961e-05
step: 520, loss: 4.055452154716477e-05
step: 530, loss: 0.00012262458039913327
/usr/local/lib/python3.7/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=0.9300766283524904, f1=0.9276879162702188,
```

```
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.322453994769603e-05
step: 10, loss: 1.8473372620064765e-05
step: 20, loss: 7.900261698523536e-05
step: 30, loss: 3.083533374592662e-05
step: 40, loss: 0.00019776266708504409
step: 50, loss: 0.03778001666069031
step: 60, loss: 0.003138007363304496
step: 70, loss: 7.850486144889146e-05
step: 80, loss: 2.2701242414768785e-05
step: 90, loss: 0.00022913914290256798
step: 100, loss: 7.194839417934418e-05
step: 110, loss: 0.001064725685864687
step: 120, loss: 9.561904153088108e-05
step: 130, loss: 3.317788286949508e-05
step: 140, loss: 7.131408347049728e-05
step: 150, loss: 6.974224379519e-05
step: 160, loss: 2.290226075274404e-05
step: 170, loss: 4.236872337060049e-05
step: 180, loss: 2.1289535652613267e-05
step: 190, loss: 4.86220778839197e-05
step: 200, loss: 0.00010820550232892856
step: 210, loss: 0.00048012894694693387
step: 220, loss: 0.00031280022813007236
step: 230, loss: 8.315958257298917e-05
step: 240, loss: 6.146874511614442e-05
step: 250, loss: 2.1878109691897407e-05
step: 260, loss: 3.05908324662596e-05
step: 270, loss: 0.000125671605928801
step: 280, loss: 0.001516978838481009
step: 290, loss: 7.100622315192595e-05
step: 300, loss: 0.0009961843024939299
step: 310, loss: 2.974529706989415e-05
step: 320, loss: 1.9989411157439463e-05
step: 330, loss: 3.060805829591118e-05
step: 340, loss: 3.1715073419036344e-05
step: 350, loss: 0.00013499293709173799
step: 360, loss: 0.0014490938046947122
step: 370, loss: 4.9644408136373386e-05
step: 380, loss: 6.469350773841143e-05
step: 390, loss: 0.0008260277681984007
step: 400, loss: 0.0003885921323671937
step: 410, loss: 0.0006077078869566321
step: 420, loss: 0.021701185032725334
step: 430, loss: 0.00038491966552101076
```

```

step: 440, loss: 0.012942506931722164
step: 450, loss: 3.945567732444033e-05
step: 460, loss: 1.7735825167619623e-05
step: 470, loss: 2.0059820599271916e-05
step: 480, loss: 5.2592607971746475e-05
step: 490, loss: 0.0007658975082449615
step: 500, loss: 3.11202384182252e-05
step: 510, loss: 0.004196592606604099
step: 520, loss: 6.894451507832855e-05
step: 530, loss: 0.0027513799723237753
/usr/local/lib/python3.7/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=0.9308235294117647, f1=0.9262564584311883,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022411728277802467
step: 10, loss: 3.510661917971447e-05
step: 20, loss: 1.561616409162525e-05
step: 30, loss: 7.006915984675288e-05
step: 40, loss: 1.9672734197229147e-05
step: 50, loss: 0.00038881489308550954
step: 60, loss: 2.0305844373069704e-05
step: 70, loss: 2.491354098310694e-05
step: 80, loss: 2.6515015633776784e-05
step: 90, loss: 5.3489504352910444e-05
step: 100, loss: 2.066740853479132e-05
step: 110, loss: 0.00014891085447743535
step: 120, loss: 1.9322593288961798e-05
step: 130, loss: 1.9046223314944655e-05
step: 140, loss: 4.1837847675196826e-05
step: 150, loss: 0.0012252581072971225
step: 160, loss: 2.0104987925151363e-05
step: 170, loss: 0.0005035370122641325
step: 180, loss: 2.9525366699090227e-05
step: 190, loss: 1.8648508557816967e-05
step: 200, loss: 1.942702874657698e-05
step: 210, loss: 3.922478572349064e-05
step: 220, loss: 2.860709537344519e-05
step: 230, loss: 2.2462279957835563e-05
step: 240, loss: 1.9602146494435146e-05
step: 250, loss: 1.9218206944060512e-05
step: 260, loss: 1.9516388420015574e-05
step: 270, loss: 1.3597163160739001e-05
step: 280, loss: 3.694225233630277e-05
step: 290, loss: 0.0001234568771906197

```

```

step: 300, loss: 0.0003701976966112852
step: 310, loss: 0.0015290880110114813
step: 320, loss: 0.0024170358665287495
step: 330, loss: 0.0026485221460461617
step: 340, loss: 0.0003120205656159669
step: 350, loss: 2.2477839593193494e-05
step: 360, loss: 3.531111724441871e-05
step: 370, loss: 0.01607969030737877
step: 380, loss: 6.243855023058131e-05
step: 390, loss: 0.00048432202311232686
step: 400, loss: 1.683429400145542e-05
step: 410, loss: 0.0005291075794957578
step: 420, loss: 1.7843718524090946e-05
step: 430, loss: 0.015256499871611595
step: 440, loss: 2.3233875253936276e-05
step: 450, loss: 1.829454231483396e-05
step: 460, loss: 0.0002767536207102239
step: 470, loss: 1.6923515431699343e-05
step: 480, loss: 2.61831701209303e-05
step: 490, loss: 4.302266097511165e-05
step: 500, loss: 2.815435072989203e-05
step: 510, loss: 2.2049371182220057e-05
step: 520, loss: 0.001436396618373692
step: 530, loss: 2.478918577253353e-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=0.9326968973747015, f1=0.922491678554446,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012039791326969862
step: 10, loss: 1.104912007576786e-05
step: 20, loss: 1.3779665096080862e-05
step: 30, loss: 2.16023745451821e-05
step: 40, loss: 3.7249621527735144e-05
step: 50, loss: 0.0008209483348764479
step: 60, loss: 3.175389065290801e-05
step: 70, loss: 1.5083448488439899e-05
step: 80, loss: 2.7286318072583526e-05
step: 90, loss: 1.677470572758466e-05
step: 100, loss: 2.894021599786356e-05
step: 110, loss: 1.8529059161664918e-05
step: 120, loss: 0.0019814774859696627
step: 130, loss: 0.02735268883407116
step: 140, loss: 0.00022560245997738093
step: 150, loss: 1.195804634335218e-05

```

```

step: 160, loss: 0.0002937245008070022
step: 170, loss: 1.5336832802859135e-05
step: 180, loss: 0.001262427307665348
step: 190, loss: 1.7988921172218397e-05
step: 200, loss: 4.332132812123746e-05
step: 210, loss: 2.277193561894819e-05
step: 220, loss: 1.7754420696292073e-05
step: 230, loss: 4.295568214729428e-05
step: 240, loss: 1.4200620171322953e-05
step: 250, loss: 6.138619937701151e-05
step: 260, loss: 1.9665047148009762e-05
step: 270, loss: 1.6182302715606056e-05
step: 280, loss: 1.7426620615879074e-05
step: 290, loss: 0.00015904277097433805
step: 300, loss: 2.4008893888094462e-05
step: 310, loss: 0.0003131819248665124
step: 320, loss: 0.0001163632477982901
step: 330, loss: 0.0002067883760901168
step: 340, loss: 0.00016816264542285353
step: 350, loss: 3.1115301680983976e-05
step: 360, loss: 0.00026693131076171994
step: 370, loss: 5.6461918575223535e-05
step: 380, loss: 1.7434018445783295e-05
step: 390, loss: 0.004918873310089111
step: 400, loss: 9.244710236089304e-05
step: 410, loss: 1.2133141353842802e-05
step: 420, loss: 5.853338007000275e-05
step: 430, loss: 2.4396331355092116e-05
step: 440, loss: 0.00011292719136690721
step: 450, loss: 6.075943019823171e-05
step: 460, loss: 5.1717775932047516e-05
step: 470, loss: 0.0024539134465157986
step: 480, loss: 1.490834529249696e-05
step: 490, loss: 1.99596252059564e-05
step: 500, loss: 1.5206322132144123e-05
step: 510, loss: 0.00013787965872325003
step: 520, loss: 5.2641044021584094e-05
step: 530, loss: 0.0001549704757053405
/usr/local/lib/python3.7/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=0.9344648750589346, f1=0.9283372365339578,
best_f1=0.9276315789473684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.4964284673624206e-05
step: 10, loss: 2.5821822418947704e-05

```

step: 20, loss: 4.0039631130639464e-05
step: 30, loss: 0.0005254324642010033
step: 40, loss: 0.04759123548865318
step: 50, loss: 0.0008022992988117039
step: 60, loss: 1.420050102751702e-05
step: 70, loss: 0.0013599920785054564
step: 80, loss: 5.173362296773121e-05
step: 90, loss: 2.4965976990642957e-05
step: 100, loss: 5.21525253134314e-05
step: 110, loss: 5.952811261522584e-05
step: 120, loss: 2.1803283743793145e-05
step: 130, loss: 0.09006313234567642
step: 140, loss: 2.1952215320197865e-05
step: 150, loss: 3.1998486520024016e-05
step: 160, loss: 2.398557808191981e-05
step: 170, loss: 1.976926068891771e-05
step: 180, loss: 1.1265178727626335e-05
step: 190, loss: 3.863612437271513e-05
step: 200, loss: 1.1704735698003788e-05
step: 210, loss: 3.552378620952368e-05
step: 220, loss: 1.4889747944835108e-05
step: 230, loss: 0.00046544079668819904
step: 240, loss: 2.074535223073326e-05
step: 250, loss: 2.4693879822734743e-05
step: 260, loss: 2.606018097139895e-05
step: 270, loss: 1.8324386473977938e-05
step: 280, loss: 1.606693331268616e-05
step: 290, loss: 1.2472133676055819e-05
step: 300, loss: 1.8614804503158666e-05
step: 310, loss: 2.5964083761209622e-05
step: 320, loss: 4.232693027006462e-05
step: 330, loss: 2.484245487721637e-05
step: 340, loss: 1.4718356396770105e-05
step: 350, loss: 0.0002504390140529722
step: 360, loss: 0.00014254823327064514
step: 370, loss: 1.2632322977879085e-05
step: 380, loss: 1.926302138599567e-05
step: 390, loss: 2.3110957044991665e-05
step: 400, loss: 2.0994955775677226e-05
step: 410, loss: 1.7720785763231106e-05
step: 420, loss: 2.493167448847089e-05
step: 430, loss: 0.000182508651050739
step: 440, loss: 4.1643634176580235e-05
step: 450, loss: 1.38280593091622e-05
step: 460, loss: 1.810451976780314e-05
step: 470, loss: 0.0010577781358733773
step: 480, loss: 1.3015953300055116e-05
step: 490, loss: 1.3924961422162596e-05

```

step: 500, loss: 3.21768966387026e-05
step: 510, loss: 1.6837791918078437e-05
step: 520, loss: 1.4666272363683674e-05
step: 530, loss: 2.682047306734603e-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=0.9339622641509434, f1=0.926829268292683,
best_f1=0.9276315789473684
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, 240.32it/s]
load_f1 = 0.9325789721829325
real_f1 = 0.9355900329102022
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 245.15it/s]

```

3 DK

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

Looking in indexes: <https://pypi.org/simple>, <https://us-python.pkg.dev/colab-wheels/public/simple/>

Collecting en_core_web_lg==2.2.5

Downloading https://github.com/explosion/spacy-models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9 MB)

| | 827.9 MB 1.0 MB/s

Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)

Requirement already satisfied: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=e046e4df233f8960c028ab11bb3345603c8a49e823de4a5ba35c65b205d8d8c0
  Stored in directory: /tmp/pip-ephem-wheel-cache-4p1_6jtf/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

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

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

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

Selected optimization level 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: 0.5716682076454163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5161290322580646, f1=0.11111111111111112,
best_f1=0.11111111111111112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5010576248168945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5185185185185186, f1=0.23529411764705882,
best_f1=0.23529411764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5724155902862549
/usr/local/lib/python3.7/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=0.4444444444444444, f1=0.358974358974359,
best_f1=0.23529411764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2615358829498291
/usr/local/lib/python3.7/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=0.43750000000000006, f1=0.42857142857142855,
best_f1=0.23529411764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2551887035369873
/usr/local/lib/python3.7/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=0.5142857142857143, f1=0.4666666666666667,
best_f1=0.23529411764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20130787789821625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5925925925925927, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.060293518006801605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.6486486486486486, f1=0.5555555555555556,
best_f1=0.5555555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030216852203011513
/usr/local/lib/python3.7/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=0.6451612903225806, f1=0.45161290322580644,
best_f1=0.5555555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002656693570315838
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6875000000000001, f1=0.48484848484848486,
best_f1=0.48484848484848486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01093687117099762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7272727272727273, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005301863886415958
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.03214917331933975
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.5454545454545454, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034033029805868864
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.5454545454545454, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003788839792832732
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.5454545454545454, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011915256269276142
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.5454545454545454, best_f1=0.5
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 144521.64it/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 = 0.5384615384615384
real_f1 = 0.6451612903225806
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, 255.37it/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: 0.6352372169494629

step: 10, loss: 0.6169479489326477

step: 20, loss: 0.36182382702827454

step: 30, loss: 0.15115930140018463


```

step: 40, loss: 0.12240631878376007
step: 50, loss: 0.07196946442127228
step: 60, loss: 0.05012175813317299
step: 70, loss: 0.22534649074077606
step: 80, loss: 0.13424672186374664
step: 90, loss: 0.19925104081630707
step: 100, loss: 0.01424767542630434
step: 110, loss: 0.24403515458106995
step: 120, loss: 0.010288301855325699
step: 130, loss: 0.031160131096839905
step: 140, loss: 0.03336016833782196
step: 150, loss: 0.12179088592529297
step: 160, loss: 0.0076835970394313335
step: 170, loss: 0.03510254621505737
step: 180, loss: 0.004093771800398827
step: 190, loss: 0.10766976326704025
step: 200, loss: 0.013249983079731464
step: 210, loss: 0.006447063758969307
step: 220, loss: 0.0849737599492073
step: 230, loss: 0.07525467872619629
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.967305524239008, f1=0.967305524239008,
best_f1=0.967305524239008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002472179476171732
step: 10, loss: 0.001087211654521525
step: 20, loss: 0.1137421503663063
step: 30, loss: 0.18779346346855164
step: 40, loss: 0.003249637084081769
step: 50, loss: 0.012441843748092651
step: 60, loss: 0.008495357818901539
step: 70, loss: 0.1706238090991974
step: 80, loss: 0.00643307575955987
step: 90, loss: 0.10310836136341095
step: 100, loss: 0.011933065950870514
step: 110, loss: 0.0928555577993393
step: 120, loss: 0.04013686999678612
step: 130, loss: 0.018045665696263313
step: 140, loss: 0.007893882691860199
step: 150, loss: 0.00495684240013361

```

```

step: 160, loss: 0.003657163120806217
step: 170, loss: 0.0012127677910029888
step: 180, loss: 0.005460716784000397
step: 190, loss: 0.03307904675602913
step: 200, loss: 0.006242942530661821
step: 210, loss: 0.002807917771860957
step: 220, loss: 0.13385145366191864
step: 230, loss: 0.1124572604894638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832402234636871, f1=0.9709821428571428,
best_f1=0.9709821428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008570033125579357
step: 10, loss: 0.0017362703802064061
step: 20, loss: 0.000939495861530304
step: 30, loss: 0.08323179185390472
step: 40, loss: 0.034239623695611954
step: 50, loss: 0.028031030669808388
step: 60, loss: 0.009038531221449375
step: 70, loss: 0.001448109745979309
step: 80, loss: 0.001309614279307425
step: 90, loss: 0.12298125773668289
step: 100, loss: 0.0013532927259802818
step: 110, loss: 0.0013619399396702647
step: 120, loss: 0.0993272215127945
step: 130, loss: 0.008354204706847668
step: 140, loss: 0.050892896950244904
step: 150, loss: 0.004152289126068354
step: 160, loss: 0.02196482941508293
step: 170, loss: 0.0017910655587911606
step: 180, loss: 0.06807699054479599
step: 190, loss: 0.006276256870478392
step: 200, loss: 0.05393047630786896
step: 210, loss: 0.011733675375580788
step: 220, loss: 0.0042769573628902435
step: 230, loss: 0.001960013760253787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9854423292273236, f1=0.9798206278026906,
best_f1=0.9798206278026906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002037501661106944
step: 10, loss: 0.00047950580483302474
step: 20, loss: 0.003242599545046687
step: 30, loss: 0.00041731807868927717
step: 40, loss: 0.0011025354033336043
step: 50, loss: 0.0071109444834291935
step: 60, loss: 0.024865737184882164
step: 70, loss: 0.0007200537947937846
step: 80, loss: 0.0007093913736753166
step: 90, loss: 0.0012817949755117297
step: 100, loss: 0.008720716461539268
step: 110, loss: 0.002647141693159938
step: 120, loss: 0.019171392545104027
step: 130, loss: 0.003346159355714917
step: 140, loss: 0.0006784514407627285
step: 150, loss: 0.0821567177772522
step: 160, loss: 0.002141886856406927
step: 170, loss: 0.0026957050431519747
step: 180, loss: 0.00038086308632045984
step: 190, loss: 0.0010535012697800994
step: 200, loss: 0.012231686152517796
step: 210, loss: 0.005476963706314564
step: 220, loss: 0.0004296615079510957
step: 230, loss: 0.001398538937792182
/usr/local/lib/python3.7/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=0.9784335981838819, f1=0.9759999999999999,
best_f1=0.9798206278026906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001085482770577073
step: 10, loss: 0.00047294949763454497
step: 20, loss: 0.16412848234176636
step: 30, loss: 0.0021191698033362627
step: 40, loss: 0.000815067847725004
step: 50, loss: 0.0014964619185775518
step: 60, loss: 0.004082332830876112
step: 70, loss: 0.0011396079789847136
```

```

step: 80, loss: 0.013534639962017536
step: 90, loss: 0.0011682870099321008
step: 100, loss: 0.0002850319433491677
step: 110, loss: 0.00023929977032821625
step: 120, loss: 0.00010252089123241603
step: 130, loss: 0.0009176433668471873
step: 140, loss: 0.008009420707821846
step: 150, loss: 0.0008159943390637636
step: 160, loss: 0.0006864324677735567
step: 170, loss: 0.0021367864683270454
step: 180, loss: 0.0005131236393935978
step: 190, loss: 0.0016509348060935736
step: 200, loss: 0.0003366968594491482
step: 210, loss: 0.0002540001296438277
step: 220, loss: 0.002136721508577466
step: 230, loss: 0.0007499762577936053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9864559819413092, f1=0.9819819819819819,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015265131369233131
step: 10, loss: 0.0009418095578439534
step: 20, loss: 0.01924920827150345
step: 30, loss: 0.0013377702562138438
step: 40, loss: 0.0004171864129602909
step: 50, loss: 0.00046013668179512024
step: 60, loss: 0.00033808802254498005
step: 70, loss: 0.001227608066983521
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.0004300783621147275
step: 90, loss: 0.0014841158408671618
step: 100, loss: 0.022482819855213165
step: 110, loss: 0.0009295637719333172
step: 120, loss: 0.0005347684491425753
step: 130, loss: 0.003151930635794997
step: 140, loss: 0.00106927624437958
step: 150, loss: 0.002066158689558506
step: 160, loss: 0.0018988043302670121
step: 170, loss: 0.0007886640378274024
step: 180, loss: 0.002959903795272112

```

```

step: 190, loss: 0.003343975404277444
step: 200, loss: 0.0006487391656264663
step: 210, loss: 0.0012008554767817259
step: 220, loss: 0.0011735325679183006
step: 230, loss: 0.008509812876582146
/usr/local/lib/python3.7/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=0.9832402234636871, f1=0.980963045912654,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012380664236843586
step: 10, loss: 0.00046888430370017886
step: 20, loss: 0.0012153706047683954
step: 30, loss: 0.0003147529496345669
step: 40, loss: 0.0001531085727037862
step: 50, loss: 0.0006672601448372006
step: 60, loss: 0.0002107533800881356
step: 70, loss: 0.0008190934313461185
step: 80, loss: 0.0001449982082704082
step: 90, loss: 0.0030171305406838655
step: 100, loss: 0.000262240442680195
step: 110, loss: 0.00026960670948028564
step: 120, loss: 0.0005611243541352451
step: 130, loss: 0.00012261133815627545
step: 140, loss: 0.000191961313248612
step: 150, loss: 0.0023442646488547325
step: 160, loss: 0.09628985077142715
step: 170, loss: 0.0018665001261979342
step: 180, loss: 0.005625266116112471
step: 190, loss: 0.0009235365432687104
step: 200, loss: 0.05981716513633728
step: 210, loss: 0.0002652766997925937
step: 220, loss: 0.0003515873395372182
step: 230, loss: 0.0015765054849907756
/usr/local/lib/python3.7/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=0.9809203142536477,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004312242672313005
step: 10, loss: 0.0002169121289625764
step: 20, loss: 0.00043284331331960857

```

```

step: 30, loss: 0.0002826303243637085
step: 40, loss: 0.007915617898106575
step: 50, loss: 0.006168997380882502
step: 60, loss: 7.248019392136484e-05
step: 70, loss: 0.00040014475234784186
step: 80, loss: 0.00017173374362755567
step: 90, loss: 5.379700451157987e-05
step: 100, loss: 0.0001277208502870053
step: 110, loss: 0.0011698628077283502
step: 120, loss: 0.00014871227904222906
step: 130, loss: 9.70148976193741e-05
step: 140, loss: 0.00023158527619671077
step: 150, loss: 0.0004331485542934388
step: 160, loss: 0.0007616067887283862
step: 170, loss: 0.0002477881498634815
step: 180, loss: 0.0012675764737650752
step: 190, loss: 0.0007174008060246706
step: 200, loss: 0.0004015429294668138
step: 210, loss: 0.00039147763163782656
step: 220, loss: 0.0007801683386787772
step: 230, loss: 0.0005254833959043026
/usr/local/lib/python3.7/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=0.9830890642615557, f1=0.9785310734463276,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.185679325833917e-05
step: 10, loss: 0.00048096096725203097
step: 20, loss: 0.0021216077730059624
step: 30, loss: 0.0001692005171207711
step: 40, loss: 0.06003952398896217
step: 50, loss: 0.00017114693764597178
step: 60, loss: 0.00018669698329176754
step: 70, loss: 0.010566574521362782
step: 80, loss: 0.0004121048841625452
step: 90, loss: 0.0003668249992188066
step: 100, loss: 0.0024336641654372215
step: 110, loss: 0.00019013545534107834
step: 120, loss: 0.00020552182104438543
step: 130, loss: 0.00018364415154792368
step: 140, loss: 0.00012961744505446404
step: 150, loss: 0.0030037562828511
step: 160, loss: 5.9392918046796694e-05
step: 170, loss: 0.00013691812637262046
step: 180, loss: 0.00022056198213249445

```

```

step: 190, loss: 0.00015396658272948116
step: 200, loss: 0.00010828254016814753
step: 210, loss: 0.0002985493920277804
step: 220, loss: 7.6774311310146e-05
step: 230, loss: 9.77688905550167e-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=0.9841986455981941, f1=0.978675645342312,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004580616659950465
step: 10, loss: 0.0003313010965939611
step: 20, loss: 0.0024640436749905348
step: 30, loss: 0.0007034221780486405
step: 40, loss: 0.0005304018268361688
step: 50, loss: 0.0001752467214828357
step: 60, loss: 0.004989480599761009
step: 70, loss: 0.0007820979808457196
step: 80, loss: 0.00013075920287519693
step: 90, loss: 0.0004158570372965187
step: 100, loss: 0.00028467527590692043
step: 110, loss: 0.00028507495881058276
step: 120, loss: 0.00037218094803392887
step: 130, loss: 0.05213937163352966
step: 140, loss: 0.04829488322138786
step: 150, loss: 0.003098799381405115
step: 160, loss: 0.0005787723348475993
step: 170, loss: 9.013717499328777e-05
step: 180, loss: 0.00028888150700367987
step: 190, loss: 0.006135253235697746
step: 200, loss: 0.00017007514543365687
step: 210, loss: 0.002980231773108244
step: 220, loss: 7.459999324055389e-05
step: 230, loss: 0.0004630677285604179
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9841628959276018, f1=0.9785310734463276,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.870665700058453e-05
step: 10, loss: 0.0010392741532996297
step: 20, loss: 0.0013533529127016664

```

```

step: 30, loss: 0.0007501942454837263
step: 40, loss: 0.0001685676397755742
step: 50, loss: 0.00024619512259960175
step: 60, loss: 0.0002381491503911093
step: 70, loss: 0.00098114472348243
step: 80, loss: 8.277566666947678e-05
step: 90, loss: 0.0001327426580246538
step: 100, loss: 0.0005240296013653278
step: 110, loss: 0.000854658370371908
step: 120, loss: 0.00021238430053927004
step: 130, loss: 0.0001553012989461422
step: 140, loss: 0.0001286633632844314
step: 150, loss: 0.05579088255763054
step: 160, loss: 8.52229495649226e-05
step: 170, loss: 0.012871512211859226
step: 180, loss: 0.0001259096898138523
step: 190, loss: 4.713390444521792e-05
step: 200, loss: 0.00019850913668051362
step: 210, loss: 0.00012078432337148115
step: 220, loss: 5.887702718609944e-05
step: 230, loss: 0.00013803226465824991
/usr/local/lib/python3.7/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=0.9841986455981941, f1=0.9796380090497738,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.577285239472985e-05
step: 10, loss: 4.8132118536159396e-05
step: 20, loss: 4.287609408493154e-05
step: 30, loss: 0.00019781995797529817
step: 40, loss: 0.0001902000658446923
step: 50, loss: 0.0032057994976639748
step: 60, loss: 0.00012643625086639076
step: 70, loss: 0.0001092143211280927
step: 80, loss: 0.0007745159673504531
step: 90, loss: 3.8491940358653665e-05
step: 100, loss: 3.937457950087264e-05
step: 110, loss: 0.0003898228460457176
step: 120, loss: 0.002394420327618718
step: 130, loss: 6.915470294188708e-05
step: 140, loss: 0.00010080420906888321
step: 150, loss: 4.7785557399038225e-05
step: 160, loss: 4.5137614506529644e-05
step: 170, loss: 0.00017551681958138943
step: 180, loss: 0.00010747456690296531

```



```

step: 190, loss: 8.293725841213018e-05
step: 200, loss: 0.0008555007516406476
step: 210, loss: 0.00017057244258467108
step: 220, loss: 0.00010236976231681183
step: 230, loss: 0.0554422028362751
/usr/local/lib/python3.7/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=0.9796839729119639, f1=0.9762174405436014,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004569226875901222
step: 10, loss: 0.0017347205430269241
step: 20, loss: 0.0002190643281210214
step: 30, loss: 0.0001096629275707528
step: 40, loss: 0.002347575267776847
step: 50, loss: 6.945618952158839e-05
step: 60, loss: 6.220788054633886e-05
step: 70, loss: 8.165084000211209e-05
step: 80, loss: 0.007823976688086987
step: 90, loss: 0.0005815388285554945
step: 100, loss: 4.020523192593828e-05
step: 110, loss: 0.03033762238919735
step: 120, loss: 0.03809726983308792
step: 130, loss: 0.0001434101432096213
step: 140, loss: 4.7845249355304986e-05
step: 150, loss: 6.449426291510463e-05
step: 160, loss: 0.0001554681221023202
step: 170, loss: 0.00030920319841243327
step: 180, loss: 0.0003839965211227536
step: 190, loss: 4.9272504838882014e-05
step: 200, loss: 0.0025110922288149595
step: 210, loss: 2.786807090160437e-05
step: 220, loss: 6.494497210951522e-05
step: 230, loss: 5.315931048244238e-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=0.9831271091113611, f1=0.9763779527559054,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.694602855830453e-05
step: 10, loss: 7.770551746943966e-05
step: 20, loss: 0.00012000366405118257

```

```

step: 30, loss: 5.333069202606566e-05
step: 40, loss: 8.011725731194019e-05
step: 50, loss: 9.670990402810276e-05
step: 60, loss: 5.2541534387273714e-05
step: 70, loss: 4.671226997743361e-05
step: 80, loss: 4.768471990246326e-05
step: 90, loss: 5.757126564276405e-05
step: 100, loss: 4.8349174903705716e-05
step: 110, loss: 4.8665417125448585e-05
step: 120, loss: 4.844251452595927e-05
step: 130, loss: 4.528999124886468e-05
step: 140, loss: 8.662914478918537e-05
step: 150, loss: 5.424210030469112e-05
step: 160, loss: 0.0009496846469119191
step: 170, loss: 2.69146948994603e-05
step: 180, loss: 6.726007268298417e-05
step: 190, loss: 4.4130763853900135e-05
step: 200, loss: 5.687946395482868e-05
step: 210, loss: 5.818430508952588e-05
step: 220, loss: 0.0001206387605634518
step: 230, loss: 0.00010402826592326164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9819819819819819, f1=0.976324689966178,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9551969419117086e-05
step: 10, loss: 2.9130438633728772e-05
step: 20, loss: 0.00010710791684687138
step: 30, loss: 7.990591257112101e-05
step: 40, loss: 0.00024079656577669084
step: 50, loss: 6.10832721577026e-05
step: 60, loss: 0.00017339947225991637
step: 70, loss: 0.002899157116189599
step: 80, loss: 3.6636691220337525e-05
step: 90, loss: 0.0001551083114463836
step: 100, loss: 4.558759246720001e-05
step: 110, loss: 0.0788603201508522
step: 120, loss: 5.8339908719062805e-05
step: 130, loss: 0.00012739196245092899
step: 140, loss: 6.372025382006541e-05
step: 150, loss: 0.0007070776191540062
step: 160, loss: 3.5712499084183946e-05
step: 170, loss: 7.675069355173036e-05
step: 180, loss: 0.0006048991926945746

```

```

step: 190, loss: 0.0003589148400351405
step: 200, loss: 0.00042109336936846375
step: 210, loss: 5.0465874664951116e-05
step: 220, loss: 8.028676529647782e-05
step: 230, loss: 3.9475249650422484e-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=0.9831271091113611, f1=0.9774774774774775,
best_f1=0.9819819819819819
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:13, 188.19it/s]
load_f1 = 0.9875141884222476
real_f1 = 0.9852104664391355
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, 233.69it/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: 0.6068850755691528
step: 10, loss: 0.5270300507545471
step: 20, loss: 0.5053499937057495
step: 30, loss: 0.16254061460494995
step: 40, loss: 0.22561100125312805
step: 50, loss: 0.10798110067844391
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.034259356558322906
step: 70, loss: 0.1549869179725647
step: 80, loss: 0.03935643285512924
step: 90, loss: 0.41250619292259216
step: 100, loss: 0.06138390302658081
step: 110, loss: 0.06954324245452881
step: 120, loss: 0.1545054316520691
step: 130, loss: 0.08995958417654037
step: 140, loss: 0.09775088727474213
step: 150, loss: 0.07525243610143661
step: 160, loss: 0.05033678933978081
step: 170, loss: 0.2547312378883362
step: 180, loss: 0.21538399159908295
step: 190, loss: 0.03015214204788208
step: 200, loss: 0.1344924420118332
step: 210, loss: 0.12014059722423553
step: 220, loss: 0.24769316613674164
step: 230, loss: 0.15683291852474213

```

```

step: 240, loss: 0.15485888719558716
step: 250, loss: 0.031582459807395935
step: 260, loss: 0.10646944493055344
step: 270, loss: 0.07084063440561295
step: 280, loss: 0.19246144592761993
step: 290, loss: 0.05080774798989296
step: 300, loss: 0.09274981915950775
step: 310, loss: 0.33367493748664856
step: 320, loss: 0.09886081516742706
step: 330, loss: 0.08466793596744537
step: 340, loss: 0.028217598795890808
step: 350, loss: 0.1929446905851364
step: 360, loss: 0.05700921639800072
step: 370, loss: 0.07332902401685715
step: 380, loss: 0.05421784520149231
step: 390, loss: 0.10769855231046677
step: 400, loss: 0.16658420860767365
step: 410, loss: 0.0550568625330925
step: 420, loss: 0.018758926540613174
step: 430, loss: 0.16032442450523376
step: 440, loss: 0.010665642097592354
step: 450, loss: 0.011556115932762623
step: 460, loss: 0.010383358225226402
step: 470, loss: 0.1633012443780899
step: 480, loss: 0.0691438615322113
step: 490, loss: 0.20907127857208252
step: 500, loss: 0.032963838428258896
step: 510, loss: 0.0762207955121994
step: 520, loss: 0.07067131251096725
step: 530, loss: 0.00677064573392272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9211009174311927, f1=0.9115895556573523,
best_f1=0.9115895556573523
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16791540384292603
step: 10, loss: 0.07842708379030228
step: 20, loss: 0.017263347283005714
step: 30, loss: 0.08232752233743668
step: 40, loss: 0.05146970972418785
step: 50, loss: 0.14575079083442688

```

step: 60, loss: 0.02370128408074379
step: 70, loss: 0.03510133549571037
step: 80, loss: 0.06345943361520767
step: 90, loss: 0.009678596630692482
step: 100, loss: 0.1275177299976349
step: 110, loss: 0.029639320448040962
step: 120, loss: 0.13946974277496338
step: 130, loss: 0.11204451322555542
step: 140, loss: 0.0481988824903965
step: 150, loss: 0.07305518537759781
step: 160, loss: 0.03767990320920944
step: 170, loss: 0.04217955842614174
step: 180, loss: 0.019640080630779266
step: 190, loss: 0.09244698286056519
step: 200, loss: 0.017504582181572914
step: 210, loss: 0.049987174570560455
step: 220, loss: 0.05584704875946045
step: 230, loss: 0.010606915690004826
step: 240, loss: 0.07583483308553696
step: 250, loss: 0.006550476420670748
step: 260, loss: 0.0016905523370951414
step: 270, loss: 0.1778893768787384
step: 280, loss: 0.03582741320133209
step: 290, loss: 0.06684666872024536
step: 300, loss: 0.10572201013565063
step: 310, loss: 0.006036738399416208
step: 320, loss: 0.049577582627534866
step: 330, loss: 0.11472567170858383
step: 340, loss: 0.022645117715001106
step: 350, loss: 0.12103116512298584
step: 360, loss: 0.2071041762828827
step: 370, loss: 0.1211489588022232
step: 380, loss: 0.03376949578523636
step: 390, loss: 0.08072736114263535
step: 400, loss: 0.05405385047197342
step: 410, loss: 0.011709876358509064
step: 420, loss: 0.08083775639533997
step: 430, loss: 0.009367204271256924
step: 440, loss: 0.07751069217920303
step: 450, loss: 0.015680793672800064
step: 460, loss: 0.017594439908862114
step: 470, loss: 0.07142972201108932
step: 480, loss: 0.2115067094564438
step: 490, loss: 0.016430621966719627
step: 500, loss: 0.28401631116867065
step: 510, loss: 0.05470903217792511
step: 520, loss: 0.029760140925645828
step: 530, loss: 0.07345228642225266

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9304229195088677, f1=0.9258079198907602,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025695525109767914
step: 10, loss: 0.06397245079278946
step: 20, loss: 0.06625008583068848
step: 30, loss: 0.026750944554805756
step: 40, loss: 0.004400196485221386
step: 50, loss: 0.15823854506015778
step: 60, loss: 0.04155378416180611
step: 70, loss: 0.0032426929101347923
step: 80, loss: 0.001478118123486638
step: 90, loss: 0.0038454090245068073
step: 100, loss: 0.044139739125967026
step: 110, loss: 0.0045501417480409145
step: 120, loss: 0.006977580487728119
step: 130, loss: 0.012887917459011078
step: 140, loss: 0.1009511724114418
step: 150, loss: 0.04071440547704697
step: 160, loss: 0.0043445550836622715
step: 170, loss: 0.1957513391971588
step: 180, loss: 0.12289343774318695
step: 190, loss: 0.021133949980139732
step: 200, loss: 0.03882087022066116
step: 210, loss: 0.06875845789909363
step: 220, loss: 0.07587747275829315
step: 230, loss: 0.11026482284069061
step: 240, loss: 0.008057557046413422
step: 250, loss: 0.027911925688385963
step: 260, loss: 0.025257764384150505
step: 270, loss: 0.0025907433591783047
step: 280, loss: 0.11655489355325699
step: 290, loss: 0.0040401979349553585
step: 300, loss: 0.070164754986763
step: 310, loss: 0.06018422171473503
step: 320, loss: 0.03054816834628582
step: 330, loss: 0.008808943443000317
step: 340, loss: 0.008652576245367527
step: 350, loss: 0.010150809772312641

```



```

step: 360, loss: 0.03257722780108452
step: 370, loss: 0.001774113392457366
step: 380, loss: 0.0030333888716995716
step: 390, loss: 0.01088614109903574
step: 400, loss: 0.005494707729667425
step: 410, loss: 0.009706443175673485
step: 420, loss: 0.09982930123806
step: 430, loss: 0.04996248707175255
step: 440, loss: 0.02025371789932251
step: 450, loss: 0.10619354248046875
step: 460, loss: 0.02138112299144268
step: 470, loss: 0.04730674624443054
step: 480, loss: 0.00990525633096695
step: 490, loss: 0.004663992673158646
step: 500, loss: 0.012781146913766861
step: 510, loss: 0.02566821128129959
step: 520, loss: 0.04856916889548302
step: 530, loss: 0.1643768548965454
/usr/local/lib/python3.7/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=0.9294449670743179, f1=0.9160739687055476,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010653638280928135
step: 10, loss: 0.012128020636737347
step: 20, loss: 0.002847106894478202
step: 30, loss: 0.017210038378834724
step: 40, loss: 0.01667102985084057
step: 50, loss: 0.002789732301607728
step: 60, loss: 0.0015232603764161468
step: 70, loss: 0.007459329906851053
step: 80, loss: 0.0017880294471979141
step: 90, loss: 0.014198895543813705
step: 100, loss: 0.01053677685558796
step: 110, loss: 0.0025070025585591793
step: 120, loss: 0.0182966236025095
step: 130, loss: 0.0032471627928316593
step: 140, loss: 0.0007368612568825483
step: 150, loss: 0.0004914907622151077
step: 160, loss: 0.009426316246390343
step: 170, loss: 0.0010272933868691325
step: 180, loss: 0.0005252013797871768
step: 190, loss: 0.023803625255823135
step: 200, loss: 0.040353596210479736
step: 210, loss: 0.051986075937747955

```

```

step: 220, loss: 0.02956167794764042
step: 230, loss: 0.010033069178462029
step: 240, loss: 0.01956423744559288
step: 250, loss: 0.05452372878789902
step: 260, loss: 0.007072998210787773
step: 270, loss: 0.014221543446183205
step: 280, loss: 0.012020662426948547
step: 290, loss: 0.04176484793424606
step: 300, loss: 0.0009094555280171335
step: 310, loss: 0.016474299132823944
step: 320, loss: 0.04027723893523216
step: 330, loss: 0.00738409161567688
step: 340, loss: 0.0171017125248909
step: 350, loss: 0.0031344592571258545
step: 360, loss: 0.009750349447131157
step: 370, loss: 0.0028635291382670403
step: 380, loss: 0.11517868936061859
step: 390, loss: 0.008960118517279625
step: 400, loss: 0.0324140228331089
step: 410, loss: 0.00035005746758542955
step: 420, loss: 0.04021234065294266
step: 430, loss: 0.09426160156726837
step: 440, loss: 0.001045728800818324
step: 450, loss: 0.0010409763781353831
step: 460, loss: 0.04382725805044174
step: 470, loss: 0.03659893572330475
step: 480, loss: 0.050328802317380905
step: 490, loss: 0.008345894515514374
step: 500, loss: 0.09038699418306351
step: 510, loss: 0.025460941717028618
step: 520, loss: 0.1461568921804428
step: 530, loss: 0.02570686861872673
/usr/local/lib/python3.7/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=0.925672594619243, f1=0.9241316270566727,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0458797812461853
step: 10, loss: 0.08434326201677322
step: 20, loss: 0.0018366037402302027
step: 30, loss: 0.04388359189033508
step: 40, loss: 0.03778635337948799
step: 50, loss: 0.0002688771055545658
step: 60, loss: 0.14304980635643005
step: 70, loss: 0.05581265687942505

```

step: 80, loss: 0.0005376214976422489
step: 90, loss: 0.006547222845256329
step: 100, loss: 0.010783782228827477
step: 110, loss: 0.0019352651434019208
step: 120, loss: 0.0007801895844750106
step: 130, loss: 0.0016171840252354741
step: 140, loss: 0.048735927790403366
step: 150, loss: 0.0003250340814702213
step: 160, loss: 0.0004917451296932995
step: 170, loss: 0.03144029900431633
step: 180, loss: 0.0011649195803329349
step: 190, loss: 0.000747202429920435
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.00031652243342250586
step: 210, loss: 0.001417039893567562
step: 220, loss: 0.047500938177108765
step: 230, loss: 0.015918642282485962
step: 240, loss: 0.001590581377968192
step: 250, loss: 0.00547062186524272
step: 260, loss: 0.0018802598351612687
step: 270, loss: 0.0006667958223260939
step: 280, loss: 0.0006794061628170311
step: 290, loss: 0.1075969710946083
step: 300, loss: 0.009569823741912842
step: 310, loss: 0.003023498458787799
step: 320, loss: 0.19117328524589539
step: 330, loss: 0.013615582138299942
step: 340, loss: 0.024121765047311783
step: 350, loss: 0.013981008902192116
step: 360, loss: 0.010986646637320518
step: 370, loss: 0.020659523084759712
step: 380, loss: 0.00021544135233853012
step: 390, loss: 0.0001454725133953616
step: 400, loss: 0.003880077041685581
step: 410, loss: 0.0018675551982596517
step: 420, loss: 0.0033034593798220158
step: 430, loss: 0.0036654879804700613
step: 440, loss: 0.032961420714855194
step: 450, loss: 0.34273192286491394
step: 460, loss: 0.061486903578042984
step: 470, loss: 0.00869095977395773
step: 480, loss: 0.001947422162629664
step: 490, loss: 0.0007800192688591778
step: 500, loss: 0.0013460369082167745
step: 510, loss: 0.0011861780658364296
step: 520, loss: 0.000893834454473108
step: 530, loss: 0.0005746575188823044
/usr/local/lib/python3.7/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=0.9296435272045028, f1=0.9294062646096307,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008166803745552897
step: 10, loss: 0.0003122281050309539
step: 20, loss: 0.10067299008369446
step: 30, loss: 0.0022010121028870344
step: 40, loss: 0.006571450736373663
step: 50, loss: 0.023062244057655334
step: 60, loss: 0.0002425971470074728
step: 70, loss: 0.09787168353796005
step: 80, loss: 0.0037571582943201065
step: 90, loss: 0.002558661624789238
step: 100, loss: 0.01312714908272028
step: 110, loss: 0.02922353520989418
step: 120, loss: 0.0002901550324168056
step: 130, loss: 0.0002784219686873257
step: 140, loss: 0.04294315725564957
step: 150, loss: 0.0006443252204917371
step: 160, loss: 0.0027552698738873005
step: 170, loss: 0.003831075271591544
step: 180, loss: 0.0006277422071434557
step: 190, loss: 0.003970210440456867
step: 200, loss: 0.003724857931956649
step: 210, loss: 0.02550944685935974
step: 220, loss: 0.003538607619702816
step: 230, loss: 0.0001106637719203718
step: 240, loss: 0.0020597209222614765
step: 250, loss: 0.0005673793493770063
step: 260, loss: 5.977265755063854e-05
step: 270, loss: 0.046080779284238815
step: 280, loss: 0.00017795883468352258
step: 290, loss: 0.00046986580127850175
step: 300, loss: 0.0007948126294650137
step: 310, loss: 0.0008116299868561327
step: 320, loss: 0.003045555204153061
step: 330, loss: 0.0005538003169931471
step: 340, loss: 0.08897415548563004
step: 350, loss: 0.0020725070498883724
step: 360, loss: 0.008904735557734966
step: 370, loss: 0.0036759013310074806
step: 380, loss: 0.00011459580855444074
step: 390, loss: 0.02380215935409069
step: 400, loss: 0.00192958430852741
```

```
step: 410, loss: 0.0016048597171902657
step: 420, loss: 0.00046669712173752487
step: 430, loss: 0.001879492192529142
step: 440, loss: 0.0028915531001985073
step: 450, loss: 0.0009463266469538212
step: 460, loss: 0.0002914041979238391
step: 470, loss: 0.0032528096344321966
step: 480, loss: 0.0016977066406980157
step: 490, loss: 0.005632099229842424
step: 500, loss: 0.0005692229606211185
step: 510, loss: 0.0020178162958472967
step: 520, loss: 0.007199910935014486
step: 530, loss: 0.004346879664808512
/usr/local/lib/python3.7/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=0.9232927970065481, f1=0.9256661991584852,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0048377676866948605
step: 10, loss: 0.00034197629429399967
step: 20, loss: 0.0001020417912513949
step: 30, loss: 6.951496470719576e-05
step: 40, loss: 0.00010909637785516679
step: 50, loss: 0.0021011861972510815
step: 60, loss: 0.0029389888513833284
step: 70, loss: 7.811714749550447e-05
step: 80, loss: 0.0006929713999852538
step: 90, loss: 0.0001264161110157147
step: 100, loss: 0.0005298318574205041
step: 110, loss: 0.008709547109901905
step: 120, loss: 0.009485532529652119
step: 130, loss: 0.0008888348238542676
step: 140, loss: 9.010423673316836e-05
step: 150, loss: 0.00048658126615919173
step: 160, loss: 0.0003971314581576735
step: 170, loss: 0.0002651326358318329
step: 180, loss: 0.0017467880388721824
step: 190, loss: 0.002867927076295018
step: 200, loss: 0.0069848294369876385
step: 210, loss: 0.001204643864184618
step: 220, loss: 6.562969065271318e-05
step: 230, loss: 0.008844178169965744
step: 240, loss: 0.03432578966021538
step: 250, loss: 0.000731959764380008
step: 260, loss: 0.00018881050345953554
```

```
step: 270, loss: 0.00016009638784453273
step: 280, loss: 0.006397663149982691
step: 290, loss: 0.002535934094339609
step: 300, loss: 0.0008035663631744683
step: 310, loss: 0.00179489073343575
step: 320, loss: 5.3138901421334594e-05
step: 330, loss: 0.061749719083309174
step: 340, loss: 0.014425262808799744
step: 350, loss: 0.0001393991697113961
step: 360, loss: 0.002048265654593706
step: 370, loss: 0.0004080454236827791
step: 380, loss: 0.0006155143491923809
step: 390, loss: 0.0006880206638015807
step: 400, loss: 0.004042471759021282
step: 410, loss: 0.0025900632608681917
step: 420, loss: 0.0019970559515058994
step: 430, loss: 0.0010673528304323554
step: 440, loss: 9.82739802566357e-05
step: 450, loss: 0.0005087197641842067
step: 460, loss: 0.0004889436531811953
step: 470, loss: 0.06522249430418015
step: 480, loss: 0.07676360011100769
step: 490, loss: 0.0005928027676418424
step: 500, loss: 0.0023549962788820267
step: 510, loss: 0.0009312143665738404
step: 520, loss: 0.02891131304204464
step: 530, loss: 0.019451601430773735
/usr/local/lib/python3.7/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=0.9251637043966324, f1=0.9216417910447762,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001332942338194698
step: 10, loss: 0.13891561329364777
step: 20, loss: 0.001000677002593875
step: 30, loss: 0.0002341052022529766
step: 40, loss: 0.0006724271806888282
step: 50, loss: 0.006406946107745171
step: 60, loss: 0.00064830482006073
step: 70, loss: 0.0007239675614982843
step: 80, loss: 0.0006823260919190943
step: 90, loss: 0.00029672510572709143
step: 100, loss: 0.00026787607930600643
step: 110, loss: 0.00037031841930001974
step: 120, loss: 0.0006973493727855384
```

```

step: 130, loss: 0.00016424534260295331
step: 140, loss: 0.0010718133999034762
step: 150, loss: 5.403541581472382e-05
step: 160, loss: 0.031012283638119698
step: 170, loss: 0.002250172197818756
step: 180, loss: 0.00033631804399192333
step: 190, loss: 0.005614744033664465
step: 200, loss: 0.0025756049435585737
step: 210, loss: 0.002515912987291813
step: 220, loss: 0.0023972413036972284
step: 230, loss: 0.002204161137342453
step: 240, loss: 0.003107290482148528
step: 250, loss: 7.390990504063666e-05
step: 260, loss: 0.0024186221417039633
step: 270, loss: 0.0007160250097513199
step: 280, loss: 0.01895330473780632
step: 290, loss: 7.950553117552772e-05
step: 300, loss: 0.0006150673725642264
step: 310, loss: 0.0004145342099945992
step: 320, loss: 0.0022815773263573647
step: 330, loss: 0.01134101115167141
step: 340, loss: 0.00044239635462872684
step: 350, loss: 6.321923865471035e-05
step: 360, loss: 0.0008671259274706244
step: 370, loss: 0.0006492460961453617
step: 380, loss: 0.00024333785404451191
step: 390, loss: 0.1720792055130005
step: 400, loss: 5.131582292960957e-05
step: 410, loss: 0.00016053838771767914
step: 420, loss: 0.007921749725937843
step: 430, loss: 0.009432761929929256
step: 440, loss: 0.004747156053781509
step: 450, loss: 0.0009881234727799892
step: 460, loss: 7.212110358523205e-05
step: 470, loss: 0.004527883604168892
step: 480, loss: 0.00010603027476463467
step: 490, loss: 0.0028193083126097918
step: 500, loss: 0.006164982449263334
step: 510, loss: 0.00030842985142953694
step: 520, loss: 0.004900705069303513
step: 530, loss: 0.0009720601956360042
/usr/local/lib/python3.7/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=0.9256198347107438, f1=0.9238532110091744,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013798772124573588
step: 10, loss: 0.00014763770741410553
step: 20, loss: 0.0020252089016139507
step: 30, loss: 0.0008707279921509326
step: 40, loss: 0.0008573809172958136
step: 50, loss: 0.0003359080001246184
step: 60, loss: 5.4514355724677444e-05
step: 70, loss: 4.062149309902452e-05
step: 80, loss: 0.0077841985039412975
step: 90, loss: 0.000613228534348309
step: 100, loss: 5.573626913246699e-05
step: 110, loss: 5.254453208181076e-05
step: 120, loss: 0.003340293886139989
step: 130, loss: 0.02368498407304287
step: 140, loss: 0.0002304536319570616
step: 150, loss: 7.28155646356754e-05
step: 160, loss: 0.0010484597878530622
step: 170, loss: 0.008016744628548622
step: 180, loss: 0.000457817135611549
step: 190, loss: 0.004272941499948502
step: 200, loss: 0.003100434085354209
step: 210, loss: 0.001289085135795176
step: 220, loss: 0.00023310797405429184
step: 230, loss: 0.010703565552830696
step: 240, loss: 0.0002621119492687285
step: 250, loss: 0.005744711495935917
step: 260, loss: 0.001359260524623096
step: 270, loss: 6.495062552858144e-05
step: 280, loss: 0.03975921869277954
step: 290, loss: 0.03164052590727806
step: 300, loss: 0.0004985590931028128
step: 310, loss: 0.0024644476361572742
step: 320, loss: 0.00031406202469952404
step: 330, loss: 0.0007957498892210424
step: 340, loss: 0.0003557655436452478
step: 350, loss: 0.000572866469155997
step: 360, loss: 0.0001207120149047114
step: 370, loss: 0.0011280316393822432
step: 380, loss: 0.0004881175409536809
step: 390, loss: 5.5821346904849634e-05
step: 400, loss: 0.0019133079331368208
step: 410, loss: 0.00028056284645572305
step: 420, loss: 0.00011238087608944625
step: 430, loss: 0.0005641643656417727
step: 440, loss: 0.011871113441884518
step: 450, loss: 9.037595737027004e-05
step: 460, loss: 6.001987640047446e-05
```



```
step: 470, loss: 3.6196750443195924e-05
step: 480, loss: 6.304238922894001e-05
step: 490, loss: 0.001381817040964961
step: 500, loss: 0.0010924231028184295
step: 510, loss: 0.00011574942618608475
step: 520, loss: 0.00030392681946977973
step: 530, loss: 0.00019418375450186431
/usr/local/lib/python3.7/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=0.9301675977653632, f1=0.9266480965645311,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004842519119847566
step: 10, loss: 3.695686609717086e-05
step: 20, loss: 3.591014683479443e-05
step: 30, loss: 7.976925553521141e-05
step: 40, loss: 7.843045023037121e-05
step: 50, loss: 6.654639582848176e-05
step: 60, loss: 3.6088858905714005e-05
step: 70, loss: 4.2513172957114875e-05
step: 80, loss: 4.473842273000628e-05
step: 90, loss: 4.2581374145811424e-05
step: 100, loss: 3.33175303239841e-05
step: 110, loss: 1.9106526451651007e-05
step: 120, loss: 8.939039253164083e-05
step: 130, loss: 3.764901703107171e-05
step: 140, loss: 0.004342029802501202
step: 150, loss: 4.606751463143155e-05
step: 160, loss: 6.03156368015334e-05
step: 170, loss: 0.0001359613670501858
step: 180, loss: 4.602349872584455e-05
step: 190, loss: 6.120264151832089e-05
step: 200, loss: 0.0004678386030718684
step: 210, loss: 0.005408049561083317
step: 220, loss: 0.00018257835472468287
step: 230, loss: 0.0004008580872323364
step: 240, loss: 0.004132514353841543
step: 250, loss: 4.549476216197945e-05
step: 260, loss: 0.0015301706735044718
step: 270, loss: 0.0008583941962569952
step: 280, loss: 0.00011715415894286707
step: 290, loss: 0.00017140436102636158
step: 300, loss: 0.00011579502461245283
step: 310, loss: 0.0032612024806439877
step: 320, loss: 0.0025749870110303164
```

```

step: 330, loss: 0.00010954911704175174
step: 340, loss: 0.0036237670574337244
step: 350, loss: 0.003109058365225792
step: 360, loss: 0.00017480950918979943
step: 370, loss: 0.004304860718548298
step: 380, loss: 0.004907242022454739
step: 390, loss: 5.680204776581377e-05
step: 400, loss: 0.000703333062119782
step: 410, loss: 0.0036529689095914364
step: 420, loss: 4.364814594737254e-05
step: 430, loss: 2.336071338504553e-05
step: 440, loss: 0.013305873610079288
step: 450, loss: 9.280774975195527e-05
step: 460, loss: 2.787525772873778e-05
step: 470, loss: 0.0001576472568558529
step: 480, loss: 3.536658186931163e-05
step: 490, loss: 6.552153354277834e-05
step: 500, loss: 0.06962060928344727
step: 510, loss: 3.50373811670579e-05
step: 520, loss: 0.0011951640481129289
step: 530, loss: 0.0004700070130638778
/usr/local/lib/python3.7/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=0.9299363057324841, f1=0.9274965800273598,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010812528198584914
step: 10, loss: 0.00017097790259867907
step: 20, loss: 4.582429755828343e-05
step: 30, loss: 0.0037738559767603874
step: 40, loss: 0.0016771790105849504
step: 50, loss: 0.0011728598037734628
step: 60, loss: 0.02078554593026638
step: 70, loss: 3.0989387596491724e-05
step: 80, loss: 0.0014103680150583386
step: 90, loss: 3.934040432795882e-05
step: 100, loss: 0.0019803803879767656
step: 110, loss: 0.0001094001199817285
step: 120, loss: 2.1248464690870605e-05
step: 130, loss: 2.2049405743018724e-05
step: 140, loss: 0.0001530126464786008
step: 150, loss: 0.007713841274380684
step: 160, loss: 0.00014721936895512044
step: 170, loss: 5.911299376748502e-05
step: 180, loss: 0.0007542589446529746

```

```

step: 190, loss: 0.00012064833572367206
step: 200, loss: 0.0027721994556486607
step: 210, loss: 0.013345113024115562
step: 220, loss: 0.0006618286133743823
step: 230, loss: 3.253415707149543e-05
step: 240, loss: 0.0002345365210203454
step: 250, loss: 4.683492807089351e-05
step: 260, loss: 3.513765841489658e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 270, loss: 3.9213948184624314e-05
step: 280, loss: 8.037065708776936e-05
step: 290, loss: 0.0005886783474124968
step: 300, loss: 8.378105121664703e-05
step: 310, loss: 5.027945735491812e-05
step: 320, loss: 0.00025599004584364593
step: 330, loss: 7.60591501602903e-05
step: 340, loss: 6.635684258071706e-05
step: 350, loss: 0.00047673817607574165
step: 360, loss: 1.9169914594385773e-05
step: 370, loss: 0.0005038821254856884
step: 380, loss: 3.497857687762007e-05
step: 390, loss: 4.9144837248604745e-05
step: 400, loss: 0.0005687687662430108
step: 410, loss: 0.00022165068367030472
step: 420, loss: 0.0058328211307525635
step: 430, loss: 4.0349365008296445e-05
step: 440, loss: 0.026302695274353027
step: 450, loss: 0.0009234407334588468
step: 460, loss: 0.0002796108019538224
step: 470, loss: 0.005059396382421255
step: 480, loss: 0.0009528544032946229
step: 490, loss: 0.0005756905884481966
step: 500, loss: 0.00021311726595740765
step: 510, loss: 0.00014937987725716084
step: 520, loss: 0.00021069913054816425
step: 530, loss: 0.0015005457680672407
/usr/local/lib/python3.7/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=0.9285384970032272, f1=0.9216589861751152,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048044483992271125
step: 10, loss: 5.602076998911798e-05
step: 20, loss: 0.00018355004431214184
step: 30, loss: 0.0008237945730797946

```

step: 40, loss: 8.718553726794198e-05
step: 50, loss: 0.017996594309806824
step: 60, loss: 0.0011184008326381445
step: 70, loss: 0.00021814402134623379
step: 80, loss: 3.601990101742558e-05
step: 90, loss: 9.081153984880075e-05
step: 100, loss: 0.0001730762596707791
step: 110, loss: 0.00011147499026264995
step: 120, loss: 4.8736932512838393e-05
step: 130, loss: 0.0003427759511396289
step: 140, loss: 8.98596626939252e-05
step: 150, loss: 0.0001500355574535206
step: 160, loss: 3.373630534042604e-05
step: 170, loss: 3.6341654777061194e-05
step: 180, loss: 6.448872591136023e-05
step: 190, loss: 6.791493069613352e-05
step: 200, loss: 4.024177178507671e-05
step: 210, loss: 0.000108292733784765
step: 220, loss: 9.320673416368663e-05
step: 230, loss: 1.619724753254559e-05
step: 240, loss: 0.0011384171666577458
step: 250, loss: 2.5259678295697086e-05
step: 260, loss: 2.5885950890369713e-05
step: 270, loss: 3.9813916373532265e-05
step: 280, loss: 0.0014016544446349144
step: 290, loss: 0.00017067418957594782
step: 300, loss: 0.00015453362721018493
step: 310, loss: 0.00035379885230213404
step: 320, loss: 6.968800880713388e-05
step: 330, loss: 0.00010328862117603421
step: 340, loss: 0.0001565238635521382
step: 350, loss: 4.828731471206993e-05
step: 360, loss: 0.07763983309268951
step: 370, loss: 0.00022111384896561503
step: 380, loss: 0.00011684570199577138
step: 390, loss: 5.475479338201694e-05
step: 400, loss: 0.06625822186470032
step: 410, loss: 0.0001793636620277539
step: 420, loss: 0.005255560390651226
step: 430, loss: 0.0064646899700164795
step: 440, loss: 0.00017178585403598845
step: 450, loss: 0.000427976599894464
step: 460, loss: 3.366605960763991e-05
step: 470, loss: 0.0003175887104589492
step: 480, loss: 6.272034079302102e-05
step: 490, loss: 0.0005458603263832629
step: 500, loss: 0.0005147717893123627
step: 510, loss: 0.0003172580327372998

```

step: 520, loss: 0.00650579109787941
step: 530, loss: 0.0012413343647494912
/usr/local/lib/python3.7/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=0.927348449791763, f1=0.9234317343173433,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017899234080687165
step: 10, loss: 2.083874096570071e-05
step: 20, loss: 1.6271658751065843e-05
step: 30, loss: 3.5913893952965736e-05
step: 40, loss: 2.8533489967230707e-05
step: 50, loss: 0.0056136962957680225
step: 60, loss: 0.001105122035369277
step: 70, loss: 3.725758506334387e-05
step: 80, loss: 1.4789171473239549e-05
step: 90, loss: 0.012110470794141293
step: 100, loss: 8.740268822293729e-05
step: 110, loss: 7.055747846607119e-05
step: 120, loss: 0.00018081540474668145
step: 130, loss: 2.30317182285944e-05
step: 140, loss: 5.5120202887337655e-05
step: 150, loss: 9.999913891078904e-05
step: 160, loss: 6.623622175538912e-05
step: 170, loss: 0.00012514185800682753
step: 180, loss: 4.207580059301108e-05
step: 190, loss: 3.7635847547790036e-05
step: 200, loss: 1.8886863472289406e-05
step: 210, loss: 3.414140883251093e-05
step: 220, loss: 4.1308954678243026e-05
step: 230, loss: 7.041620119707659e-05
step: 240, loss: 3.892889799317345e-05
step: 250, loss: 5.189314833842218e-05
step: 260, loss: 5.12800324941054e-05
step: 270, loss: 1.715080907160882e-05
step: 280, loss: 2.447043516440317e-05
step: 290, loss: 2.0007864804938436e-05
step: 300, loss: 2.761822906904854e-05
step: 310, loss: 0.002617995487526059
step: 320, loss: 0.003935040440410376
step: 330, loss: 0.000463979464257136
step: 340, loss: 3.2771607948234305e-05
step: 350, loss: 5.0793438276741654e-05
step: 360, loss: 6.575473526027054e-05
step: 370, loss: 0.0010346139315515757

```

```

step: 380, loss: 0.0001727267081150785
step: 390, loss: 0.00015245167014654726
step: 400, loss: 7.866944361012429e-05
step: 410, loss: 0.0002911369956564158
step: 420, loss: 0.00023052378674037755
step: 430, loss: 0.002605936722829938
step: 440, loss: 2.014218807744328e-05
step: 450, loss: 2.4038556148298085e-05
step: 460, loss: 0.0008706369553692639
step: 470, loss: 4.605111098499037e-05
step: 480, loss: 1.5396215530927293e-05
step: 490, loss: 1.966912896023132e-05
step: 500, loss: 4.792063555214554e-05
step: 510, loss: 8.509409963153303e-05
step: 520, loss: 0.0003612112195696682
step: 530, loss: 3.373314393684268e-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=0.9237248479176415, f1=0.9255813953488372,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013433685526251793
step: 10, loss: 1.2684370631177444e-05
step: 20, loss: 2.535588100727182e-05
step: 30, loss: 3.102606933680363e-05
step: 40, loss: 4.70008235424757e-05
step: 50, loss: 8.919546962715685e-05
step: 60, loss: 2.245123869215604e-05
step: 70, loss: 1.9143157260259613e-05
step: 80, loss: 0.00036911392817273736
step: 90, loss: 7.55640139686875e-05
step: 100, loss: 0.0001212373172165826
step: 110, loss: 1.677835280133877e-05
step: 120, loss: 5.31489058630541e-05
step: 130, loss: 0.01809215545654297
step: 140, loss: 0.0005278319586068392
step: 150, loss: 1.3451803170028143e-05
step: 160, loss: 0.0001636508823139593
step: 170, loss: 0.02335994690656662
step: 180, loss: 2.2607842765864916e-05
step: 190, loss: 9.902994497679174e-05
step: 200, loss: 3.610688145272434e-05
step: 210, loss: 1.706901275611017e-05
step: 220, loss: 2.559877793828491e-05
step: 230, loss: 0.00021064931934233755

```

```

step: 240, loss: 8.055589569266886e-05
step: 250, loss: 1.9315173631184734e-05
step: 260, loss: 0.0011424832046031952
step: 270, loss: 4.021845597890206e-05
step: 280, loss: 2.722992212511599e-05
step: 290, loss: 0.0022940943017601967
step: 300, loss: 1.9385986888664775e-05
step: 310, loss: 0.00012624975352082402
step: 320, loss: 0.0004607446026057005
step: 330, loss: 3.68961482308805e-05
step: 340, loss: 5.001311728847213e-05
step: 350, loss: 2.8243997803656384e-05
step: 360, loss: 8.216247078962624e-05
step: 370, loss: 4.013631769339554e-05
step: 380, loss: 2.8735190426232293e-05
step: 390, loss: 0.010970995761454105
step: 400, loss: 9.603436046745628e-05
step: 410, loss: 0.0005105241434648633
step: 420, loss: 0.004399530123919249
step: 430, loss: 0.00017565286543685943
step: 440, loss: 0.0015468198107555509
step: 450, loss: 0.00038095080526545644
step: 460, loss: 0.00028465516516007483
step: 470, loss: 1.8365311916568317e-05
step: 480, loss: 2.973545269924216e-05
step: 490, loss: 2.0689038137788884e-05
step: 500, loss: 1.3541184671339579e-05
step: 510, loss: 8.366075780941173e-05
step: 520, loss: 4.324637848185375e-05
step: 530, loss: 3.015454058186151e-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=0.9280373831775701, f1=0.9288040949278734,
best_f1=0.9258079198907602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.825344043027144e-05
step: 10, loss: 2.106198553519789e-05
step: 20, loss: 2.1095480406074785e-05
step: 30, loss: 1.669642369961366e-05
step: 40, loss: 0.011480805464088917
step: 50, loss: 3.4396038245176896e-05
step: 60, loss: 1.4110956726653967e-05
step: 70, loss: 0.0007053720764815807
step: 80, loss: 2.8868636945844628e-05
step: 90, loss: 1.6074296581791714e-05

```

```
step: 100, loss: 5.708055687136948e-05
step: 110, loss: 0.0005345553508959711
step: 120, loss: 0.00015766896831337363
step: 130, loss: 0.1225668266415596
step: 140, loss: 3.524774365359917e-05
step: 150, loss: 5.9591937315417454e-05
step: 160, loss: 4.1004524973686785e-05
step: 170, loss: 1.466994854126824e-05
step: 180, loss: 1.3824294910591561e-05
step: 190, loss: 3.2191190257435665e-05
step: 200, loss: 2.060011502180714e-05
step: 210, loss: 0.00019804674957413226
step: 220, loss: 2.3378393962047994e-05
step: 230, loss: 0.00035688249045051634
step: 240, loss: 1.9046845409320667e-05
step: 250, loss: 5.75024496356491e-05
step: 260, loss: 2.2675207219435833e-05
step: 270, loss: 1.8514287148718722e-05
step: 280, loss: 2.7203752324567176e-05
step: 290, loss: 1.7388860214850865e-05
step: 300, loss: 1.6308928024955094e-05
step: 310, loss: 0.00011050837929360569
step: 320, loss: 2.435659189359285e-05
step: 330, loss: 5.3668543841922656e-05
step: 340, loss: 1.1075168004026636e-05
step: 350, loss: 9.152971870207693e-06
step: 360, loss: 0.0008773290901444852
step: 370, loss: 1.0825596291397233e-05
step: 380, loss: 9.166065865429118e-05
step: 390, loss: 0.0007199097890406847
step: 400, loss: 2.0920715542160906e-05
step: 410, loss: 5.2917865104973316e-05
step: 420, loss: 5.8021956647280604e-05
step: 430, loss: 2.5841651222435758e-05
step: 440, loss: 0.04216241091489792
step: 450, loss: 1.2937648534716573e-05
step: 460, loss: 2.5341592845506966e-05
step: 470, loss: 0.0011697913287207484
step: 480, loss: 1.0575998203421477e-05
step: 490, loss: 1.6074372979346663e-05
step: 500, loss: 2.687212145247031e-05
step: 510, loss: 2.9788381652906537e-05
step: 520, loss: 1.681190224189777e-05
step: 530, loss: 4.385184001876041e-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=0.9299953639313862, f1=0.9283402681460935,
```



```

best_f1=0.9258079198907602
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:23, 242.78it/s]
load_f1 = 0.9308291342189647
real_f1 = 0.929784304726939
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 243.14it/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: 0.5754161477088928
step: 10, loss: 0.34897416830062866
step: 20, loss: 0.33680588006973267
step: 30, loss: 0.3245941698551178
step: 40, loss: 0.17857354879379272
step: 50, loss: 0.42970961332321167
step: 60, loss: 0.21534954011440277
step: 70, loss: 0.1345774233341217
step: 80, loss: 0.19068831205368042
step: 90, loss: 0.3903895616531372
step: 100, loss: 0.3830837607383728
step: 110, loss: 0.21810723841190338
step: 120, loss: 0.2634891867637634
step: 130, loss: 0.18180352449417114
step: 140, loss: 0.22463566064834595
step: 150, loss: 0.2992844581604004
step: 160, loss: 0.4282751977443695
step: 170, loss: 0.24396869540214539
step: 180, loss: 0.17517004907131195
step: 190, loss: 0.23118151724338531
step: 200, loss: 0.2480921447277069
step: 210, loss: 0.22242961823940277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5847457627118645, f1=0.6184210526315789,
best_f1=0.6184210526315789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09215374290943146
step: 10, loss: 0.19424279034137726
step: 20, loss: 0.13589885830879211
step: 30, loss: 0.25408855080604553
step: 40, loss: 0.16608019173145294
step: 50, loss: 0.1865537464618683
step: 60, loss: 0.3740105628967285
step: 70, loss: 0.1019446924328804
step: 80, loss: 0.19353920221328735
step: 90, loss: 0.0729399099946022
step: 100, loss: 0.02645980753004551
step: 110, loss: 0.1336207240819931
step: 120, loss: 0.19244356453418732
step: 130, loss: 0.045261938124895096
step: 140, loss: 0.17426897585391998
step: 150, loss: 0.2178621143102646
step: 160, loss: 0.19809725880622864
step: 170, loss: 0.15896113216876984
step: 180, loss: 0.19714535772800446
step: 190, loss: 0.2915029227733612
step: 200, loss: 0.06213611364364624
step: 210, loss: 0.2059468775987625
/usr/local/lib/python3.7/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=0.5791666666666667, f1=0.6180257510729614,
best_f1=0.6184210526315789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.056916821748018265
step: 10, loss: 0.22601334750652313
step: 20, loss: 0.18865181505680084
step: 30, loss: 0.1775895059108734
step: 40, loss: 0.2587548494338989
step: 50, loss: 0.0677310973405838
step: 60, loss: 0.17310650646686554
step: 70, loss: 0.18380656838417053
step: 80, loss: 0.15639671683311462
step: 90, loss: 0.06564126163721085
step: 100, loss: 0.38172706961631775
step: 110, loss: 0.1758250892162323
step: 120, loss: 0.06315157562494278
step: 130, loss: 0.08544506877660751
step: 140, loss: 0.08082298189401627
step: 150, loss: 0.2948904037475586

```

```

step: 160, loss: 0.04193758964538574
step: 170, loss: 0.14176353812217712
step: 180, loss: 0.07614223659038544
step: 190, loss: 0.17220734059810638
step: 200, loss: 0.04450476914644241
step: 210, loss: 0.060477763414382935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6174242424242424, f1=0.6460348162475823,
best_f1=0.6460348162475823
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07751989364624023
step: 10, loss: 0.05858113616704941
step: 20, loss: 0.10046353936195374
step: 30, loss: 0.12040171027183533
step: 40, loss: 0.043440911918878555
step: 50, loss: 0.11289485543966293
step: 60, loss: 0.08644522726535797
step: 70, loss: 0.23739352822303772
step: 80, loss: 0.0478249229490757
step: 90, loss: 0.019197765737771988
step: 100, loss: 0.22565466165542603
step: 110, loss: 0.0758756548166275
step: 120, loss: 0.15291297435760498
step: 130, loss: 0.1585957109928131
step: 140, loss: 0.12072647362947464
step: 150, loss: 0.015219754539430141
step: 160, loss: 0.062286268919706345
step: 170, loss: 0.12730471789836884
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.30224016308784485
step: 190, loss: 0.06246887147426605
step: 200, loss: 0.15138830244541168
step: 210, loss: 0.22620746493339539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6191369606003753, f1=0.6262230919765167,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031495027244091034
step: 10, loss: 0.04400403052568436
step: 20, loss: 0.1092258021235466
step: 30, loss: 0.1526562124490738
step: 40, loss: 0.11098208278417587
step: 50, loss: 0.038952622562646866
step: 60, loss: 0.12124894559383392
step: 70, loss: 0.03506694361567497
step: 80, loss: 0.01877041533589363
step: 90, loss: 0.018823983147740364
step: 100, loss: 0.007247685454785824
step: 110, loss: 0.37933269143104553
step: 120, loss: 0.03304028511047363
step: 130, loss: 0.09857113659381866
step: 140, loss: 0.13645945489406586
step: 150, loss: 0.02339921146631241
step: 160, loss: 0.22565779089927673
step: 170, loss: 0.09083732962608337
step: 180, loss: 0.10891518741846085
step: 190, loss: 0.02743968740105629
step: 200, loss: 0.032890912145376205
step: 210, loss: 0.026693886145949364
/usr/local/lib/python3.7/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=0.5989304812834224, f1=0.6088560885608856,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041234590113162994
step: 10, loss: 0.07757211476564407
step: 20, loss: 0.011473987251520157
step: 30, loss: 0.0008761457866057754
step: 40, loss: 0.12912943959236145
step: 50, loss: 0.04195917770266533
step: 60, loss: 0.11894697695970535
step: 70, loss: 0.013871383853256702
step: 80, loss: 0.020208412781357765
step: 90, loss: 0.2789342701435089
step: 100, loss: 0.0019679698161780834
step: 110, loss: 0.004630249459296465
step: 120, loss: 0.029633989557623863

```

```

step: 130, loss: 0.05263344570994377
step: 140, loss: 0.06180809438228607
step: 150, loss: 0.017792511731386185
step: 160, loss: 0.0008801759104244411
step: 170, loss: 0.044599805027246475
step: 180, loss: 0.01816042698919773
step: 190, loss: 0.07421299070119858
step: 200, loss: 0.002232134807854891
step: 210, loss: 0.02166329324245453
/usr/local/lib/python3.7/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=0.6022304832713754, f1=0.6266924564796905,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01778092421591282
step: 10, loss: 0.027529697865247726
step: 20, loss: 0.02742069959640503
step: 30, loss: 0.008053366094827652
step: 40, loss: 0.007625663187354803
step: 50, loss: 0.07642576098442078
step: 60, loss: 0.025988861918449402
step: 70, loss: 0.01455730851739645
step: 80, loss: 0.015382027253508568
step: 90, loss: 0.042977672070264816
step: 100, loss: 0.01636553555727005
step: 110, loss: 0.07840146124362946
step: 120, loss: 0.06441865116357803
step: 130, loss: 0.030124252662062645
step: 140, loss: 0.02217555232346058
step: 150, loss: 0.01294983271509409
step: 160, loss: 0.022096781060099602
step: 170, loss: 0.0027846277225762606
step: 180, loss: 0.05177310109138489
step: 190, loss: 0.07682502269744873
step: 200, loss: 0.04156903177499771
step: 210, loss: 0.03357374295592308
/usr/local/lib/python3.7/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=0.6045627376425856, f1=0.6141414141414142,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006667039822787046

```

```

step: 10, loss: 0.029011180624365807
step: 20, loss: 0.013170499354600906
step: 30, loss: 0.02815677411854267
step: 40, loss: 0.005652984604239464
step: 50, loss: 0.0015820886474102736
step: 60, loss: 0.005193675868213177
step: 70, loss: 0.019417177885770798
step: 80, loss: 0.045830339193344116
step: 90, loss: 0.0076947519555687904
step: 100, loss: 0.06707482039928436
step: 110, loss: 0.05549147352576256
step: 120, loss: 0.0009799775434657931
step: 130, loss: 0.004160101059824228
step: 140, loss: 0.09188925474882126
step: 150, loss: 0.0057669151574373245
step: 160, loss: 0.01292396429926157
step: 170, loss: 0.02808372862637043
step: 180, loss: 0.04501677304506302
step: 190, loss: 0.006997307762503624
step: 200, loss: 0.02116977982223034
step: 210, loss: 0.12573958933353424
/usr/local/lib/python3.7/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=0.5927505330490406, f1=0.6065934065934065,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003554431488737464
step: 10, loss: 0.012477933429181576
step: 20, loss: 0.0016197956865653396
step: 30, loss: 0.008464448153972626
step: 40, loss: 0.039383772760629654
step: 50, loss: 0.018687374889850616
step: 60, loss: 0.05996128171682358
step: 70, loss: 0.009101674892008305
step: 80, loss: 0.001036471570841968
step: 90, loss: 0.0008243973716162145
step: 100, loss: 0.09330246597528458
step: 110, loss: 0.005596187897026539
step: 120, loss: 0.00044861395144835114
step: 130, loss: 0.01002300065010786
step: 140, loss: 0.013138187117874622
step: 150, loss: 0.05883076786994934
step: 160, loss: 0.0007739128195680678
step: 170, loss: 0.0028359624557197094
step: 180, loss: 0.08091127872467041

```



```

step: 190, loss: 0.0075018578208982944
step: 200, loss: 0.07800713181495667
step: 210, loss: 0.033624082803726196
/usr/local/lib/python3.7/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=0.6019417475728155, f1=0.6309278350515464,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005543468054383993
step: 10, loss: 0.0011802536901086569
step: 20, loss: 0.00048385930131189525
step: 30, loss: 0.0037716059014201164
step: 40, loss: 0.0004392936534713954
step: 50, loss: 0.02900778315961361
step: 60, loss: 0.017684824764728546
step: 70, loss: 0.01462011132389307
step: 80, loss: 0.07530470937490463
step: 90, loss: 0.011998825706541538
step: 100, loss: 0.004915957804769278
step: 110, loss: 0.03378044068813324
step: 120, loss: 0.0009534343262203038
step: 130, loss: 0.059244442731142044
step: 140, loss: 0.0016663004644215107
step: 150, loss: 0.01010469812899828
step: 160, loss: 0.0047928160056471825
step: 170, loss: 0.0003921282186638564
step: 180, loss: 0.03247610479593277
step: 190, loss: 0.054492752999067307
step: 200, loss: 0.00022022011398803443
step: 210, loss: 0.0030661243945360184
/usr/local/lib/python3.7/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=0.5777777777777777, f1=0.6027397260273972,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004382758866995573
step: 10, loss: 0.027394382283091545
step: 20, loss: 0.0016063520452007651
step: 30, loss: 0.00017900655802804977
step: 40, loss: 0.0012579687172546983
step: 50, loss: 0.007555163465440273
step: 60, loss: 0.017456725239753723

```

```

step: 70, loss: 0.00013615679927170277
step: 80, loss: 0.04921211302280426
step: 90, loss: 0.00018276690389029682
step: 100, loss: 0.005201069638133049
step: 110, loss: 0.029693452641367912
step: 120, loss: 0.012608510442078114
step: 130, loss: 0.007048106752336025
step: 140, loss: 0.00015561898180749267
step: 150, loss: 7.553531031589955e-05
step: 160, loss: 0.003952209372073412
step: 170, loss: 0.00019789945508819073
step: 180, loss: 0.003115114290267229
step: 190, loss: 0.003241482423618436
step: 200, loss: 0.0007256928947754204
step: 210, loss: 0.0009230420691892505
/usr/local/lib/python3.7/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=0.5786163522012578, f1=0.5990990990990991,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004898762563243508
step: 10, loss: 0.00016454202705062926
step: 20, loss: 0.0024673324078321457
step: 30, loss: 0.010574220679700375
step: 40, loss: 0.0035417149774730206
step: 50, loss: 0.0005476083606481552
step: 60, loss: 0.010624601505696774
step: 70, loss: 0.0212835855782032
step: 80, loss: 0.00013003953790757805
step: 90, loss: 0.0033166627399623394
step: 100, loss: 0.0019397067371755838
step: 110, loss: 0.02236640453338623
step: 120, loss: 8.423945837421343e-05
step: 130, loss: 0.00011465246643638238
step: 140, loss: 0.00030877956305630505
step: 150, loss: 0.005382104776799679
step: 160, loss: 0.003842402948066592
step: 170, loss: 0.0029839512426406145
step: 180, loss: 0.0002803936367854476
step: 190, loss: 0.00024134409613907337
step: 200, loss: 0.016038792207837105
step: 210, loss: 0.0026203731540590525
/usr/local/lib/python3.7/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=0.5780885780885782, f1=0.5915492957746479,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003959659021347761
step: 10, loss: 0.0004168015730101615
step: 20, loss: 0.029290398582816124
step: 30, loss: 0.0020008222199976444
step: 40, loss: 0.030802922323346138
step: 50, loss: 0.0002691170375328511
step: 60, loss: 0.0075044576078653336
step: 70, loss: 0.035782698541879654
step: 80, loss: 0.02288677543401718
step: 90, loss: 0.00018215205636806786
step: 100, loss: 0.0006781459087505937
step: 110, loss: 0.0062377252615988255
step: 120, loss: 0.00047855020966380835
step: 130, loss: 0.00013661070261150599
step: 140, loss: 0.00019852496916428208
step: 150, loss: 0.0009397142566740513
step: 160, loss: 0.0008228300721384585
step: 170, loss: 0.0002426288992865011
step: 180, loss: 0.05647513270378113
step: 190, loss: 0.0003672963357530534
step: 200, loss: 0.00026617434923537076
step: 210, loss: 0.007963738404214382
/usr/local/lib/python3.7/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=0.6109979633401222, f1=0.6177105831533477,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.459472933784127e-05
step: 10, loss: 0.0009100593160837889
step: 20, loss: 0.0041825734078884125
step: 30, loss: 0.030371377244591713
step: 40, loss: 0.00012392886856105179
step: 50, loss: 0.00023335765581578016
step: 60, loss: 0.00023585671442560852
step: 70, loss: 9.679255890659988e-05
step: 80, loss: 0.0001394123100908473
step: 90, loss: 0.00024812135961838067
step: 100, loss: 0.004431779496371746
step: 110, loss: 0.00018474430544301867
step: 120, loss: 0.030117614194750786

```

```

step: 130, loss: 0.0048146797344088554
step: 140, loss: 0.004255276173353195
step: 150, loss: 0.0423605851829052
step: 160, loss: 0.0008775480091571808
step: 170, loss: 0.0003646161640062928
step: 180, loss: 0.0230998732149601
step: 190, loss: 0.0011648783693090081
step: 200, loss: 0.00017647934146225452
step: 210, loss: 0.011745554395020008
/usr/local/lib/python3.7/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=0.5995717344753747, f1=0.5995525727069352,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031964152003638446
step: 10, loss: 7.538275531260297e-05
step: 20, loss: 0.0004654311924241483
step: 30, loss: 0.0007594944909214973
step: 40, loss: 0.0017785230884328485
step: 50, loss: 9.939592564478517e-05
step: 60, loss: 0.0006121420301496983
step: 70, loss: 9.164350194623694e-05
step: 80, loss: 0.00010517945338506252
step: 90, loss: 0.0059972116723656654
step: 100, loss: 0.00020162825239822268
step: 110, loss: 9.634912566980347e-05
step: 120, loss: 0.00017317943274974823
step: 130, loss: 0.0004923117230646312
step: 140, loss: 0.0005465620197355747
step: 150, loss: 0.0002745723177213222
step: 160, loss: 0.0008386206463910639
step: 170, loss: 0.00012958265142515302
step: 180, loss: 7.778783765388653e-05
step: 190, loss: 0.000346519926097244
step: 200, loss: 0.0011219721054658294
step: 210, loss: 0.0002865456335712224
/usr/local/lib/python3.7/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=0.5906040268456376, f1=0.5903890160183066,
best_f1=0.6262230919765167
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, 449.30it/s]
load_f1 = 0.6177024482109228
real_f1 = 0.6162570888468809
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, 237.89it/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 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: 0.5781556367874146
step: 10, loss: 0.3651178181171417
step: 20, loss: 0.28146082162857056
step: 30, loss: 0.4162902534008026
step: 40, loss: 0.431736022233963
step: 50, loss: 0.3078685402870178
step: 60, loss: 0.27681249380111694
step: 70, loss: 0.2715517580509186
step: 80, loss: 0.21888846158981323
step: 90, loss: 0.27903980016708374
step: 100, loss: 0.3007155656814575
step: 110, loss: 0.41351085901260376
step: 120, loss: 0.13273294270038605
step: 130, loss: 0.09032072126865387
step: 140, loss: 0.04428458586335182
step: 150, loss: 0.2377743273973465
step: 160, loss: 0.11071816086769104
step: 170, loss: 0.24011974036693573
step: 180, loss: 0.0329873226583004
step: 190, loss: 0.23725582659244537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6648199445983378, f1=0.6894736842105263,
best_f1=0.6894736842105263
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3292226493358612
step: 10, loss: 0.030503004789352417
step: 20, loss: 0.04860392585396767
step: 30, loss: 0.13702496886253357
step: 40, loss: 0.18070170283317566

```

```

step: 50, loss: 0.08396454155445099
step: 60, loss: 0.3262006342411041
step: 70, loss: 0.1379903256893158
step: 80, loss: 0.2805653214454651
step: 90, loss: 0.18668873608112335
step: 100, loss: 0.03578969091176987
step: 110, loss: 0.19949667155742645
step: 120, loss: 0.17020179331302643
step: 130, loss: 0.1846172958612442
step: 140, loss: 0.08279148489236832
step: 150, loss: 0.08512203395366669
step: 160, loss: 0.18918465077877045
step: 170, loss: 0.09154900908470154
step: 180, loss: 0.1347181648015976
step: 190, loss: 0.047605015337467194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7783505154639175, f1=0.7712082262210797,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052005548030138016
step: 10, loss: 0.024145202711224556
step: 20, loss: 0.0537530817091465
step: 30, loss: 0.019120702520012856
step: 40, loss: 0.034008197486400604
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.12317796796560287
step: 60, loss: 0.048151031136512756
step: 70, loss: 0.0979607030749321
step: 80, loss: 0.0904306173324585
step: 90, loss: 0.04781348630785942
step: 100, loss: 0.04356329143047333
step: 110, loss: 0.010820053517818451
step: 120, loss: 0.005885446444153786
step: 130, loss: 0.037091080099344254
step: 140, loss: 0.013133429922163486
step: 150, loss: 0.10760357230901718
step: 160, loss: 0.014717970974743366
step: 170, loss: 0.1325584352016449
step: 180, loss: 0.04984697327017784
step: 190, loss: 0.02680937573313713

```



```

/usr/local/lib/python3.7/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=0.7611548556430445, f1=0.7633587786259541,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01646447740495205
step: 10, loss: 0.09173670411109924
step: 20, loss: 0.005806737579405308
step: 30, loss: 0.01804433949291706
step: 40, loss: 0.008216137066483498
step: 50, loss: 0.02476007118821144
step: 60, loss: 0.00913719367235899
step: 70, loss: 0.08946962654590607
step: 80, loss: 0.0492778941988945
step: 90, loss: 0.018214883282780647
step: 100, loss: 0.028090786188840866
step: 110, loss: 0.0037251010071486235
step: 120, loss: 0.07813743501901627
step: 130, loss: 0.10438932478427887
step: 140, loss: 0.03410527482628822
step: 150, loss: 0.003733607940375805
step: 160, loss: 0.019943170249462128
step: 170, loss: 0.05929122865200043
step: 180, loss: 0.0039019808173179626
step: 190, loss: 0.10178987681865692
/usr/local/lib/python3.7/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=0.765498652291105, f1=0.7362637362637363,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059490397572517395
step: 10, loss: 0.014555852860212326
step: 20, loss: 0.07736384868621826
step: 30, loss: 0.028599750250577927
step: 40, loss: 0.06379853188991547
step: 50, loss: 0.11058677732944489
step: 60, loss: 0.10603503882884979
step: 70, loss: 0.009313062764704227
step: 80, loss: 0.06134568527340889
step: 90, loss: 0.014005156233906746
step: 100, loss: 0.0019133922178298235
step: 110, loss: 0.006024503614753485

```

```

step: 120, loss: 0.001712458673864603
step: 130, loss: 0.09378519654273987
step: 140, loss: 0.04349592328071594
step: 150, loss: 0.00976741686463356
step: 160, loss: 0.09520883858203888
step: 170, loss: 0.0018833102658390999
step: 180, loss: 0.016767732799053192
step: 190, loss: 0.05887294188141823
/usr/local/lib/python3.7/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=0.7613636363636365, f1=0.7392550143266476,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031315740197896957
step: 10, loss: 0.11751872301101685
step: 20, loss: 0.010304464027285576
step: 30, loss: 0.001194456941448152
step: 40, loss: 0.04937964305281639
step: 50, loss: 0.008120400831103325
step: 60, loss: 0.02148006483912468
step: 70, loss: 0.023270834237337112
step: 80, loss: 0.07851868122816086
step: 90, loss: 0.020009370520710945
step: 100, loss: 0.0007778799626976252
step: 110, loss: 0.015138719230890274
step: 120, loss: 0.015792282298207283
step: 130, loss: 0.006239823065698147
step: 140, loss: 0.14041130244731903
step: 150, loss: 0.004379370249807835
step: 160, loss: 0.041440851986408234
step: 170, loss: 0.024154407903552055
step: 180, loss: 0.0036286129616200924
step: 190, loss: 0.06533271819353104
/usr/local/lib/python3.7/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=0.7539267015706806, f1=0.753315649867374,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05278968811035156
step: 10, loss: 0.00590901542454958
step: 20, loss: 0.24157099425792694
step: 30, loss: 0.05090751126408577

```

```

step: 40, loss: 0.003931001760065556
step: 50, loss: 0.005446234252303839
step: 60, loss: 0.004884723573923111
step: 70, loss: 0.004603864625096321
step: 80, loss: 0.0015509696677327156
step: 90, loss: 0.01160832867026329
step: 100, loss: 0.0007699996931478381
step: 110, loss: 0.001050099148415029
step: 120, loss: 0.0013658558018505573
step: 130, loss: 0.0011198635911569
step: 140, loss: 0.0007293605012819171
step: 150, loss: 0.07394613325595856
step: 160, loss: 0.010570647194981575
step: 170, loss: 0.001692900201305747
step: 180, loss: 0.009756483137607574
step: 190, loss: 0.004276285879313946
/usr/local/lib/python3.7/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=0.7183462532299743, f1=0.733509234828496,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008616011589765549
step: 10, loss: 0.08413570374250412
step: 20, loss: 0.00390825979411602
step: 30, loss: 0.047203682363033295
step: 40, loss: 0.0013579040532931685
step: 50, loss: 0.0329391248524189
step: 60, loss: 0.0018450086936354637
step: 70, loss: 0.0014761112397536635
step: 80, loss: 0.00020237949502188712
step: 90, loss: 0.0003920479502994567
step: 100, loss: 0.054959218949079514
step: 110, loss: 0.00021702238882426172
step: 120, loss: 0.0005322650540620089
step: 130, loss: 0.008772765286266804
step: 140, loss: 0.0033537978306412697
step: 150, loss: 0.0014638499123975635
step: 160, loss: 0.0013429500395432115
step: 170, loss: 0.001319325529038906
step: 180, loss: 0.0027845248114317656
step: 190, loss: 0.008395574055612087
/usr/local/lib/python3.7/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=0.7368421052631579, f1=0.7179487179487181,

```

```

best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017652444075793028
step: 10, loss: 0.01109409797936678
step: 20, loss: 0.0012315706117078662
step: 30, loss: 0.0009347114246338606
step: 40, loss: 0.0030479049310088158
step: 50, loss: 0.0008420681115239859
step: 60, loss: 0.0008006941643543541
step: 70, loss: 0.00028883470804430544
step: 80, loss: 0.001711062854155898
step: 90, loss: 0.08329276740550995
step: 100, loss: 0.049923595041036606
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.12441044300794601
step: 120, loss: 0.0607636421918869
step: 130, loss: 0.0012600296176970005
step: 140, loss: 0.002371434587985277
step: 150, loss: 0.0012854676460847259
step: 160, loss: 0.0011021955870091915
step: 170, loss: 0.001114308601245284
step: 180, loss: 0.0008783533703535795
step: 190, loss: 0.002409471897408366
/usr/local/lib/python3.7/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=0.7182320441988949, f1=0.7329545454545454,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005781440995633602
step: 10, loss: 0.0005287314997985959
step: 20, loss: 0.12259991466999054
step: 30, loss: 0.006094001699239016
step: 40, loss: 0.08862647414207458
step: 50, loss: 0.005250099115073681
step: 60, loss: 0.003388690995052457
step: 70, loss: 0.004794489126652479
step: 80, loss: 0.007285118568688631
step: 90, loss: 0.018144501373171806
step: 100, loss: 0.004628406371921301
step: 110, loss: 0.009938178583979607
step: 120, loss: 0.10293714702129364
step: 130, loss: 0.0025090528652071953
step: 140, loss: 0.0033427115995436907

```

```

step: 150, loss: 0.048482827842235565
step: 160, loss: 0.002523761475458741
step: 170, loss: 0.0015887905610725284
step: 180, loss: 0.007825538516044617
step: 190, loss: 0.0035434970632195473
/usr/local/lib/python3.7/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=0.7258485639686684, f1=0.7547169811320755,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008330558193847537
step: 10, loss: 0.0005510500050149858
step: 20, loss: 0.0037446317728608847
step: 30, loss: 0.007028399966657162
step: 40, loss: 0.000465084973257035
step: 50, loss: 0.004665793385356665
step: 60, loss: 0.00042071426287293434
step: 70, loss: 0.0013454153668135405
step: 80, loss: 0.0007249650661833584
step: 90, loss: 0.0001522720413049683
step: 100, loss: 0.000383655889891088
step: 110, loss: 0.00035442167427390814
step: 120, loss: 0.00045174072147347033
step: 130, loss: 0.009693591855466366
step: 140, loss: 0.00995846837759018
step: 150, loss: 0.00044161194819025695
step: 160, loss: 0.003744250861927867
step: 170, loss: 0.00028439785819500685
step: 180, loss: 0.018313312903046608
step: 190, loss: 0.0006600628839805722
/usr/local/lib/python3.7/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=0.7478753541076487, f1=0.7449856733524355,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003610410203691572
step: 10, loss: 0.0034788944758474827
step: 20, loss: 0.00018645540694706142
step: 30, loss: 0.0005321408971212804
step: 40, loss: 0.0021056567784398794
step: 50, loss: 0.012632220983505249
step: 60, loss: 0.0005879594245925546

```

```

step: 70, loss: 0.00042256363667547703
step: 80, loss: 0.0002565158065408468
step: 90, loss: 0.0014621085720136762
step: 100, loss: 0.00047146997530944645
step: 110, loss: 0.00021389291214291006
step: 120, loss: 0.05905042588710785
step: 130, loss: 0.00027294931351207197
step: 140, loss: 0.0008567156037315726
step: 150, loss: 0.0004883937654085457
step: 160, loss: 0.00010200437827734277
step: 170, loss: 0.0010208056773990393
step: 180, loss: 8.1390913692303e-05
step: 190, loss: 0.00019491332932375371
/usr/local/lib/python3.7/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=0.7592067988668556, f1=0.735042735042735,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006749593652784824
step: 10, loss: 0.0003496795252431184
step: 20, loss: 0.000160784344188869
step: 30, loss: 0.00039345870027318597
step: 40, loss: 0.00018011126667261124
step: 50, loss: 0.0004739888245239854
step: 60, loss: 0.0002035990182776004
step: 70, loss: 0.0007342970347963274
step: 80, loss: 0.0009894571267068386
step: 90, loss: 0.00042712187860161066
step: 100, loss: 0.00026153656654059887
step: 110, loss: 0.001549556152895093
step: 120, loss: 0.04072147607803345
step: 130, loss: 0.0005265639629215002
step: 140, loss: 0.0004876283055637032
step: 150, loss: 0.00013315830437932163
step: 160, loss: 0.00018781019025482237
step: 170, loss: 0.0037873408291488886
step: 180, loss: 0.00015828147297725081
step: 190, loss: 0.013553522527217865
/usr/local/lib/python3.7/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=0.7627118644067796, f1=0.7282913165266107,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002899389364756644
step: 10, loss: 0.0001610146282473579
step: 20, loss: 0.0006373528158292174
step: 30, loss: 7.826575892977417e-05
step: 40, loss: 0.001036704401485622
step: 50, loss: 0.0008241728064604104
step: 60, loss: 0.0017631346127018332
step: 70, loss: 0.00017791874415706843
step: 80, loss: 0.000181359049747698
step: 90, loss: 0.0012186556123197079
step: 100, loss: 0.0016059176996350288
step: 110, loss: 0.0013683594297617674
step: 120, loss: 0.0030848965980112553
step: 130, loss: 0.00039884535362944007
step: 140, loss: 0.0008437593351118267
step: 150, loss: 0.00015430942585226148
step: 160, loss: 0.00024468559422530234
step: 170, loss: 0.0002353420277358964
step: 180, loss: 0.0003337294328957796
step: 190, loss: 0.0006480647716671228
/usr/local/lib/python3.7/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=0.7605633802816902, f1=0.7247191011235956,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008156168041750789
step: 10, loss: 0.00012554162822198123
step: 20, loss: 0.0011248444207012653
step: 30, loss: 0.00012077545397914946
step: 40, loss: 0.0005397142376750708
step: 50, loss: 0.00024067745835054666
step: 60, loss: 0.000339469377649948
step: 70, loss: 0.003619471797719598
step: 80, loss: 0.00018069089855998755
step: 90, loss: 0.000760513823479414
step: 100, loss: 0.0005421423120424151
step: 110, loss: 0.0010003320639953017
step: 120, loss: 0.00015704514225944877
step: 130, loss: 0.0001575939531903714
step: 140, loss: 0.0024228240363299847
step: 150, loss: 0.0005206016357988119
step: 160, loss: 0.0021069336216896772
step: 170, loss: 0.00028730789199471474
step: 180, loss: 0.00042099205893464386

```

```

step: 190, loss: 0.0015341655816882849
/usr/local/lib/python3.7/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=0.7584269662921349, f1=0.7211267605633802,
best_f1=0.7712082262210797
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:13, 154.29it/s]
load_f1 = 0.5317647058823529
real_f1 = 0.5181818181818182
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:25, 174.44it/s]

```


3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Textual/Abt-Buy \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm 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: 0.6306821703910828
step: 10, loss: 0.36284351348876953
step: 20, loss: 0.2973402738571167
step: 30, loss: 0.3935077488422394
step: 40, loss: 0.2766757607460022
step: 50, loss: 0.2982383966445923
step: 60, loss: 0.21576091647148132
step: 70, loss: 0.3728560209274292
step: 80, loss: 0.41314560174942017
step: 90, loss: 0.23155570030212402
step: 100, loss: 0.1938527375459671
step: 110, loss: 0.20932899415493011
step: 120, loss: 0.21340259909629822
step: 130, loss: 0.05852949991822243
step: 140, loss: 0.20626647770404816
step: 150, loss: 0.4107530117034912
step: 160, loss: 0.23180192708969116
step: 170, loss: 0.2268071174621582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7082405345211581, f1=0.7085201793721974,
best_f1=0.7085201793721974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10224848240613937
step: 10, loss: 0.27777716517448425
step: 20, loss: 0.0765891969203949
step: 30, loss: 0.20884814858436584
step: 40, loss: 0.058085449039936066
step: 50, loss: 0.16556879878044128
step: 60, loss: 0.12990914285182953
step: 70, loss: 0.2709573209285736
step: 80, loss: 0.08118011057376862
step: 90, loss: 0.20788989961147308
step: 100, loss: 0.1983616203069687
step: 110, loss: 0.04128403589129448
step: 120, loss: 0.061320990324020386
step: 130, loss: 0.07783524692058563
step: 140, loss: 0.2827879786491394
step: 150, loss: 0.18462884426116943
step: 160, loss: 0.14389805495738983
step: 170, loss: 0.046896085143089294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7785888077858881, f1=0.7599067599067598,
best_f1=0.7599067599067598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09545227140188217
step: 10, loss: 0.06240960955619812
step: 20, loss: 0.018247684463858604
step: 30, loss: 0.12008510529994965
step: 40, loss: 0.07988769561052322
step: 50, loss: 0.22565878927707672
step: 60, loss: 0.08249008655548096
step: 70, loss: 0.06648249179124832
step: 80, loss: 0.01876227930188179
step: 90, loss: 0.17157304286956787
step: 100, loss: 0.06653354316949844

```

```

step: 110, loss: 0.12331406772136688
step: 120, loss: 0.04109431430697441
step: 130, loss: 0.14459358155727386
step: 140, loss: 0.10505212098360062
step: 150, loss: 0.024299517273902893
step: 160, loss: 0.06170161813497543
step: 170, loss: 0.06355420500040054
/usr/local/lib/python3.7/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=0.7748691099476439, f1=0.7858942065491185,
best_f1=0.7599067599067598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028475504368543625
step: 10, loss: 0.022709421813488007
step: 20, loss: 0.006555144675076008
step: 30, loss: 0.09322801232337952
step: 40, loss: 0.009798322804272175
step: 50, loss: 0.10953555256128311
step: 60, loss: 0.1721712201833725
step: 70, loss: 0.025228669866919518
step: 80, loss: 0.13080261647701263
step: 90, loss: 0.0846557542681694
step: 100, loss: 0.17322812974452972
step: 110, loss: 0.13749441504478455
step: 120, loss: 0.007440519984811544
step: 130, loss: 0.15530642867088318
step: 140, loss: 0.02412419579923153
step: 150, loss: 0.07127035409212112
step: 160, loss: 0.20335786044597626
step: 170, loss: 0.014884996227920055
/usr/local/lib/python3.7/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=0.7535545023696684, f1=0.7806004618937643,
best_f1=0.7599067599067598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004570798482745886
step: 10, loss: 0.013168823905289173
step: 20, loss: 0.04260359704494476
step: 30, loss: 0.022016990929841995
step: 40, loss: 0.12827959656715393
step: 50, loss: 0.030243681743741035
step: 60, loss: 0.052696455270051956

```

```

step: 70, loss: 0.1772533357143402
step: 80, loss: 0.14135012030601501
step: 90, loss: 0.07802389562129974
step: 100, loss: 0.006185742560774088
step: 110, loss: 0.05487363040447235
step: 120, loss: 0.018313728272914886
step: 130, loss: 0.05006563290953636
step: 140, loss: 0.00669903913512826
step: 150, loss: 0.008853505365550518
step: 160, loss: 0.10006801038980484
step: 170, loss: 0.004428825341165066
/usr/local/lib/python3.7/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=0.7599067599067598, f1=0.7703016241299303,
best_f1=0.7599067599067598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018657641485333443
step: 10, loss: 0.05683773383498192
step: 20, loss: 0.02720525860786438
step: 30, loss: 0.002854445483535528
step: 40, loss: 0.014826199039816856
step: 50, loss: 0.11003416776657104
step: 60, loss: 0.01069454662501812
step: 70, loss: 0.06922110170125961
step: 80, loss: 0.016742577776312828
step: 90, loss: 0.028059568256139755
step: 100, loss: 0.005805283784866333
step: 110, loss: 0.0012525838101282716
step: 120, loss: 0.02566388249397278
step: 130, loss: 0.009006829932332039
step: 140, loss: 0.0435485914349556
step: 150, loss: 0.010161536745727062
step: 160, loss: 0.029279692098498344
step: 170, loss: 0.010162964463233948
/usr/local/lib/python3.7/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=0.7648456057007125, f1=0.7754137115839244,
best_f1=0.7599067599067598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004746209364384413
step: 10, loss: 0.003224150277674198
step: 20, loss: 0.003505786182358861

```

```

step: 30, loss: 0.047513436526060104
step: 40, loss: 0.0076171099208295345
step: 50, loss: 0.012057202868163586
step: 60, loss: 0.019150542095303535
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.002416432835161686
step: 80, loss: 0.014028949663043022
step: 90, loss: 0.0010464690858498216
step: 100, loss: 0.0005819108919240534
step: 110, loss: 0.03462501987814903
step: 120, loss: 0.0010721449507400393
step: 130, loss: 0.1338406503200531
step: 140, loss: 0.0014807148836553097
step: 150, loss: 0.005183515138924122
step: 160, loss: 0.0008088255999609828
step: 170, loss: 0.12175340205430984
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7790432801822323, f1=0.7661469933184855,
best_f1=0.7661469933184855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011850345879793167
step: 10, loss: 0.006974825169891119
step: 20, loss: 0.0034685914870351553
step: 30, loss: 0.01989583484828472
step: 40, loss: 0.00017785344971343875
step: 50, loss: 0.006331956945359707
step: 60, loss: 0.037461020052433014
step: 70, loss: 0.0012885028263553977
step: 80, loss: 0.004040556028485298
step: 90, loss: 0.002369260648265481
step: 100, loss: 0.03629009425640106
step: 110, loss: 0.20171275734901428
step: 120, loss: 0.008407055400311947
step: 130, loss: 0.00473770359531045
step: 140, loss: 0.0011068639578297734
step: 150, loss: 0.0006869998760521412
step: 160, loss: 0.005761581938713789
step: 170, loss: 0.0029939545784145594
/usr/local/lib/python3.7/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=0.7612293144208038, f1=0.7788018433179723,
best_f1=0.7661469933184855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005375835578888655
step: 10, loss: 0.005033374764025211
step: 20, loss: 0.00037251971662044525
step: 30, loss: 0.021238502115011215
step: 40, loss: 0.06410954892635345
step: 50, loss: 0.0037274002097547054
step: 60, loss: 0.06191441789269447
step: 70, loss: 0.0019179085502400994
step: 80, loss: 0.018140295520424843
step: 90, loss: 0.02110704965889454
step: 100, loss: 0.036127910017967224
step: 110, loss: 0.00311398901976645
step: 120, loss: 0.01890787109732628
step: 130, loss: 0.0653165876865387
step: 140, loss: 0.007220447529107332
step: 150, loss: 0.0026820374187082052
step: 160, loss: 0.0046689631417393684
step: 170, loss: 0.0120914988219738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7912621359223301, f1=0.7990430622009569,
best_f1=0.7990430622009569
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06365052610635757
step: 10, loss: 0.0023306452203541994
step: 20, loss: 0.08312075585126877
step: 30, loss: 0.10536537319421768
step: 40, loss: 0.013754351064562798
step: 50, loss: 0.037607431411743164
step: 60, loss: 0.000747558893635869
step: 70, loss: 0.007871626876294613
step: 80, loss: 0.0026254872791469097
step: 90, loss: 0.015232667326927185
step: 100, loss: 0.011220937594771385
step: 110, loss: 0.000496559776365757

```

```

step: 120, loss: 0.0006974849966354668
step: 130, loss: 0.0009944585617631674
step: 140, loss: 0.018836528062820435
step: 150, loss: 0.0014501443365588784
step: 160, loss: 0.0014997334219515324
step: 170, loss: 0.00020127395691815764
/usr/local/lib/python3.7/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=0.7486910994764399, f1=0.7930174563591021,
best_f1=0.7990430622009569
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003632934531196952
step: 10, loss: 0.0329350009560585
step: 20, loss: 0.003093196079134941
step: 30, loss: 0.00014792087313253433
step: 40, loss: 0.00020305249199736863
step: 50, loss: 0.01242704689502716
step: 60, loss: 0.0006511559477075934
step: 70, loss: 0.0002469451283104718
step: 80, loss: 0.03485646843910217
step: 90, loss: 0.0004414098511915654
step: 100, loss: 0.0029781947378069162
step: 110, loss: 0.012367538176476955
step: 120, loss: 0.00030079481075517833
step: 130, loss: 0.00013183719420339912
step: 140, loss: 0.025953449308872223
step: 150, loss: 0.0020551211200654507
step: 160, loss: 0.01160750724375248
step: 170, loss: 0.0024422556161880493
/usr/local/lib/python3.7/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=0.7819905213270142, f1=0.7934272300469484,
best_f1=0.7990430622009569
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022231239825487137
step: 10, loss: 0.10298644006252289
step: 20, loss: 0.0034245920833200216
step: 30, loss: 0.0005009055021218956
step: 40, loss: 0.0006723743863403797
step: 50, loss: 0.0015613138675689697
step: 60, loss: 0.0007323637255467474
step: 70, loss: 0.04461103677749634

```



```

step: 80, loss: 0.0003845604369416833
step: 90, loss: 0.0032001868821680546
step: 100, loss: 0.00717270839959383
step: 110, loss: 0.0007061440264806151
step: 120, loss: 0.01171985361725092
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.012254010885953903
step: 140, loss: 0.012196743860840797
step: 150, loss: 0.01108810305595398
step: 160, loss: 0.0004246599564794451
step: 170, loss: 0.0002698799653444439
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.7953488372093023, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05741036310791969
step: 10, loss: 0.0005670576356351376
step: 20, loss: 0.0015414867084473372
step: 30, loss: 0.00013321560982149094
step: 40, loss: 0.0007321593584492803
step: 50, loss: 0.007814242504537106
step: 60, loss: 0.004003442823886871
step: 70, loss: 0.001742254476994276
step: 80, loss: 0.003642813768237829
step: 90, loss: 0.00011200350127182901
step: 100, loss: 0.00025312029174529016
step: 110, loss: 0.00045954110100865364
step: 120, loss: 0.015145749785006046
step: 130, loss: 0.00021304877009242773
step: 140, loss: 0.0003473362303338945
step: 150, loss: 0.008103648200631142
step: 160, loss: 0.0006111040711402893
step: 170, loss: 0.021289803087711334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7990196078431373, f1=0.8019323671497586,

```

```

best_f1=0.8019323671497586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001815431984141469
step: 10, loss: 0.00021886274043936282
step: 20, loss: 0.0002425064885756001
step: 30, loss: 0.00027048320043832064
step: 40, loss: 0.00020604619930963963
step: 50, loss: 0.008441982790827751
step: 60, loss: 0.00032945573912002146
step: 70, loss: 0.0006503229378722608
step: 80, loss: 0.0011323073413223028
step: 90, loss: 0.0009523491025902331
step: 100, loss: 0.00015644611266907305
step: 110, loss: 0.00013899248733650893
step: 120, loss: 0.024413563311100006
step: 130, loss: 0.001796963857486844
step: 140, loss: 0.00013484405644703656
step: 150, loss: 0.0008002772810868919
step: 160, loss: 0.00016170386516023427
step: 170, loss: 0.0003732974291779101
/usr/local/lib/python3.7/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=0.7865707434052758, f1=0.7999999999999999,
best_f1=0.8019323671497586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017131647327914834
step: 10, loss: 0.0001945019030245021
step: 20, loss: 0.035363588482141495
step: 30, loss: 0.0005622311728075147
step: 40, loss: 0.0014299058821052313
step: 50, loss: 0.00021138113515917212
step: 60, loss: 0.004660416394472122
step: 70, loss: 0.00015328841982409358
step: 80, loss: 0.0005605347687378526
step: 90, loss: 0.002429481130093336
step: 100, loss: 0.0003127149830106646
step: 110, loss: 0.000126956743770279
step: 120, loss: 0.0002279706677654758
step: 130, loss: 0.00026218523271381855
step: 140, loss: 0.00015536211139988154
step: 150, loss: 0.0005007992149330676
step: 160, loss: 0.0002115977549692616
step: 170, loss: 0.00024877916439436376

```

```

/usr/local/lib/python3.7/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=0.7892156862745098, f1=0.8057553956834533,
best_f1=0.8019323671497586
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, 214.64it/s]
load_f1 = 0.6081632653061224
real_f1 = 0.6085192697768763
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:25, 173.94it/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: 0.6120129227638245
step: 10, loss: 0.621228814125061
step: 20, loss: 0.4405178129673004
step: 30, loss: 0.22840973734855652
step: 40, loss: 0.227983757853508
step: 50, loss: 0.0506976880133152
step: 60, loss: 0.058961909264326096
step: 70, loss: 0.11547849327325821
step: 80, loss: 0.03747252747416496
step: 90, loss: 0.16254684329032898
step: 100, loss: 0.008816612884402275
step: 110, loss: 0.25396621227264404
step: 120, loss: 0.01188000850379467
step: 130, loss: 0.040173642337322235
step: 140, loss: 0.04387220740318298
step: 150, loss: 0.040450818836688995
step: 160, loss: 0.13102063536643982
step: 170, loss: 0.15929284691810608
step: 180, loss: 0.04305034503340721
step: 190, loss: 0.030837014317512512
step: 200, loss: 0.149497851729393
step: 210, loss: 0.02424868382513523
step: 220, loss: 0.009606434032320976

```

```

step: 230, loss: 0.026999756693840027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9573770491803277, f1=0.9578713968957872,
best_f1=0.9578713968957872
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020124122500419617
step: 10, loss: 0.025628870353102684
step: 20, loss: 0.18921953439712524
step: 30, loss: 0.15221524238586426
step: 40, loss: 0.08972673863172531
step: 50, loss: 0.008576219901442528
step: 60, loss: 0.009662940166890621
step: 70, loss: 0.0010515454923734069
step: 80, loss: 0.09213929623365402
step: 90, loss: 0.008797656744718552
step: 100, loss: 0.15480025112628937
step: 110, loss: 0.12422103434801102
step: 120, loss: 0.11887349933385849
step: 130, loss: 0.0066624777391552925
step: 140, loss: 0.002849721582606435
step: 150, loss: 0.004394428804516792
step: 160, loss: 0.07540426403284073
step: 170, loss: 0.0017878683283925056
step: 180, loss: 0.08102727681398392
step: 190, loss: 0.013884554617106915
step: 200, loss: 0.0066092489287257195
step: 210, loss: 0.002187700942158699
step: 220, loss: 0.03703606501221657
step: 230, loss: 0.06294190138578415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9648924122310306, f1=0.9727891156462585,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015432857908308506
step: 10, loss: 0.005322849377989769
step: 20, loss: 0.009454818442463875
step: 30, loss: 0.00891720037907362
step: 40, loss: 0.0414101667702198
step: 50, loss: 0.009945322759449482
step: 60, loss: 0.016662729904055595
step: 70, loss: 0.0026916866190731525
step: 80, loss: 0.0025856790598481894
step: 90, loss: 0.018379464745521545
step: 100, loss: 0.0005423406255431473
step: 110, loss: 0.0006570170517079532
step: 120, loss: 0.005061608739197254
step: 130, loss: 0.0007404709467664361
step: 140, loss: 0.01446995697915554
step: 150, loss: 0.006027133669704199
step: 160, loss: 0.06572186201810837
step: 170, loss: 0.0220989678055048
step: 180, loss: 0.028623143211007118
step: 190, loss: 0.009920659475028515
step: 200, loss: 0.004947133362293243
step: 210, loss: 0.09021814167499542
step: 220, loss: 0.002277111168950796
step: 230, loss: 0.008570586331188679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9711111111111111, f1=0.9688888888888889,
best_f1=0.9688888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003128174867015332
step: 10, loss: 0.0008814004249870777
step: 20, loss: 0.0014638920547440648
step: 30, loss: 0.0015647634863853455
step: 40, loss: 0.0018144322093576193
step: 50, loss: 0.00039254582952708006
step: 60, loss: 0.0033676442690193653
step: 70, loss: 0.002553738886490464
step: 80, loss: 0.0024229097180068493
step: 90, loss: 0.0024755820631980896
step: 100, loss: 0.004753496963530779

```

```

step: 110, loss: 0.0016515961615368724
step: 120, loss: 0.0031697533559054136
step: 130, loss: 0.0024250366259366274
step: 140, loss: 0.00070545932976529
step: 150, loss: 0.10661470144987106
step: 160, loss: 0.10743238776922226
step: 170, loss: 0.007509829010814428
step: 180, loss: 0.003530848305672407
step: 190, loss: 0.004375654272735119
step: 200, loss: 0.002177590737119317
step: 210, loss: 0.03617164492607117
step: 220, loss: 0.0005616039852611721
step: 230, loss: 0.006030102726072073
/usr/local/lib/python3.7/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=0.9683257918552037, f1=0.971815107102593,
best_f1=0.9688888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031601432710886
step: 10, loss: 0.0007484462694264948
step: 20, loss: 0.06988026201725006
step: 30, loss: 0.014145351946353912
step: 40, loss: 0.001450452720746398
step: 50, loss: 0.0022391248494386673
step: 60, loss: 0.0008865951676853001
step: 70, loss: 0.0020204265601933002
step: 80, loss: 0.0005720174522139132
step: 90, loss: 0.0010997939389199018
step: 100, loss: 0.00020483025582507253
step: 110, loss: 0.00020233314717188478
step: 120, loss: 9.735061757965013e-05
step: 130, loss: 0.002236417029052973
step: 140, loss: 0.0022365874610841274
step: 150, loss: 0.003968460951000452
step: 160, loss: 0.016722505912184715
step: 170, loss: 0.015748126432299614
step: 180, loss: 0.0026299594901502132
step: 190, loss: 0.07534866780042648
step: 200, loss: 0.013050246052443981
step: 210, loss: 0.02088979259133339
step: 220, loss: 0.0006139914621599019
step: 230, loss: 0.0037941972259432077
/usr/local/lib/python3.7/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=0.9707865168539327, f1=0.9718785151856018,
best_f1=0.9688888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027853590436279774
step: 10, loss: 0.001272656605578959
step: 20, loss: 0.0010601339163258672
step: 30, loss: 0.0004089047433808446
step: 40, loss: 0.00022880033066030592
step: 50, loss: 0.0024288659915328026
step: 60, loss: 0.00018790891044773161
step: 70, loss: 0.0006412736838683486
step: 80, loss: 0.0005934145883657038
step: 90, loss: 0.005353877786546946
step: 100, loss: 0.007597532123327255
step: 110, loss: 0.00072244624607265
step: 120, loss: 0.00033337585045956075
step: 130, loss: 0.0003062129835598171
step: 140, loss: 0.0006191994762048125
step: 150, loss: 0.010220308788120747
step: 160, loss: 0.0017253314144909382
step: 170, loss: 0.0006297912332229316
step: 180, loss: 0.020365282893180847
step: 190, loss: 0.014664997346699238
step: 200, loss: 0.0005238482845015824
step: 210, loss: 0.0004899489576928318
step: 220, loss: 0.00025281761190854013
step: 230, loss: 0.024851134046912193
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9731543624161074, f1=0.9700332963374029,
best_f1=0.9700332963374029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004334199707955122
step: 10, loss: 0.0002057678357232362
step: 20, loss: 0.0010083595989271998
step: 30, loss: 0.00031564326491206884
step: 40, loss: 0.0007534394389949739
step: 50, loss: 0.00026829898706637323
step: 60, loss: 0.0010782700264826417

```

```

step: 70, loss: 0.03942576050758362
step: 80, loss: 0.06732559204101562
step: 90, loss: 0.001104413764551282
step: 100, loss: 0.14960098266601562
step: 110, loss: 0.00039901304990053177
step: 120, loss: 0.0006186665268614888
step: 130, loss: 0.00022834964329376817
step: 140, loss: 0.0021142156329005957
step: 150, loss: 0.0006785074947401881
step: 160, loss: 0.07733098417520523
step: 170, loss: 0.001345776254311204
step: 180, loss: 0.0024844282306730747
step: 190, loss: 0.00036947568878531456
step: 200, loss: 0.016160128638148308
step: 210, loss: 0.00019470539700705558
step: 220, loss: 0.0009778202511370182
step: 230, loss: 0.0007076504989527166
/usr/local/lib/python3.7/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=0.9719416386083053, f1=0.972129319955407,
best_f1=0.9700332963374029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023148389300331473
step: 10, loss: 0.0006266393465921283
step: 20, loss: 0.00021239783382043242
step: 30, loss: 0.0002032161719398573
step: 40, loss: 0.0019259699620306492
step: 50, loss: 0.0002967270265799016
step: 60, loss: 0.0007740239379927516
step: 70, loss: 0.02282172441482544
step: 80, loss: 0.0005114414962008595
step: 90, loss: 0.0001647252938710153
step: 100, loss: 0.0008732809801585972
step: 110, loss: 0.02113417536020279
step: 120, loss: 0.004625893663614988
step: 130, loss: 0.007793961558490992
step: 140, loss: 0.00019770061771851033
step: 150, loss: 0.00015257333870977163
step: 160, loss: 0.0001396393054164946
step: 170, loss: 0.0003534733841661364
step: 180, loss: 0.030885107815265656
step: 190, loss: 0.0005147689953446388
step: 200, loss: 0.0012387807946652174
step: 210, loss: 0.00036136762355454266
step: 220, loss: 0.04323078691959381

```

```

step: 230, loss: 0.00016517704352736473
/usr/local/lib/python3.7/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=0.9721913236929923, f1=0.9678848283499446,
best_f1=0.9700332963374029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015306135173887014
step: 10, loss: 0.00023315080034080893
step: 20, loss: 0.0015472613740712404
step: 30, loss: 7.799669401720166e-05
step: 40, loss: 0.07457059621810913
step: 50, loss: 0.00017675806884653866
step: 60, loss: 0.00027276729815639555
step: 70, loss: 0.00027789882733486593
step: 80, loss: 0.0012540628667920828
step: 90, loss: 0.0008973304647952318
step: 100, loss: 0.00016988179413601756
step: 110, loss: 0.0001234780065715313
step: 120, loss: 0.0005167751805856824
step: 130, loss: 0.00014640139124821872
step: 140, loss: 9.57196025410667e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 150, loss: 0.00890036765486002
step: 160, loss: 8.852846804074943e-05
step: 170, loss: 0.0007345370249822736
step: 180, loss: 0.008379649370908737
step: 190, loss: 0.00021152749832253903
step: 200, loss: 0.15122286975383759
step: 210, loss: 0.02961094118654728
step: 220, loss: 0.0004090812581125647
step: 230, loss: 0.0008820408256724477
/usr/local/lib/python3.7/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=0.9696969696969697, f1=0.972129319955407,
best_f1=0.9700332963374029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002383348997682333
step: 10, loss: 0.00013760995352640748
step: 20, loss: 0.0008131799986585975
step: 30, loss: 0.00019101840734947473
step: 40, loss: 0.0003319120442029089
step: 50, loss: 0.00016415806021541357

```

```

step: 60, loss: 0.0266561359167099
step: 70, loss: 0.0008864377159625292
step: 80, loss: 0.00016513647278770804
step: 90, loss: 0.00019295979291200638
step: 100, loss: 8.51886288728565e-05
step: 110, loss: 0.00010285902681061998
step: 120, loss: 0.00014329739497043192
step: 130, loss: 0.0001972253230633214
step: 140, loss: 0.036194272339344025
step: 150, loss: 0.01982288807630539
step: 160, loss: 8.171053923433647e-05
step: 170, loss: 5.53015161131043e-05
step: 180, loss: 8.93485193955712e-05
step: 190, loss: 0.00587886106222868
step: 200, loss: 6.872640369692817e-05
step: 210, loss: 0.00018338541849516332
step: 220, loss: 0.00015917923883534968
step: 230, loss: 0.029521355405449867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753914988814317, f1=0.9742441209406495,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019111167057417333
step: 10, loss: 0.00046766424202360213
step: 20, loss: 0.0008151134243234992
step: 30, loss: 0.005923447664827108
step: 40, loss: 0.00018541907775215805
step: 50, loss: 0.00012922289897687733
step: 60, loss: 0.00031561427749693394
step: 70, loss: 0.0018834611400961876
step: 80, loss: 0.00013504078378900886
step: 90, loss: 9.76806550170295e-05
step: 100, loss: 0.00023641531879547983
step: 110, loss: 0.00567863741889596
step: 120, loss: 0.00039643043419346213
step: 130, loss: 7.568900764454156e-05
step: 140, loss: 0.0001726011687424034
step: 150, loss: 0.016936028376221657
step: 160, loss: 0.0002248376840725541
step: 170, loss: 0.0035407179966568947

```

```

step: 180, loss: 0.0038786442019045353
step: 190, loss: 0.0002283173962496221
step: 200, loss: 0.0006870842189528048
step: 210, loss: 5.193382821744308e-05
step: 220, loss: 8.61346343299374e-05
step: 230, loss: 9.220497304340824e-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=0.9754464285714286, 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: 0.00011702313349815086
step: 10, loss: 6.703726103296503e-05
step: 20, loss: 0.00017264894268009812
step: 30, loss: 0.00028005754575133324
step: 40, loss: 0.00016858644085004926
step: 50, loss: 0.0028785422910004854
step: 60, loss: 0.00018527822976466268
step: 70, loss: 7.805446512065828e-05
step: 80, loss: 6.643019150942564e-05
step: 90, loss: 0.00037274573696777225
step: 100, loss: 0.0003862883895635605
step: 110, loss: 6.54870891594328e-05
step: 120, loss: 5.754351877840236e-05
step: 130, loss: 7.889376138336957e-05
step: 140, loss: 0.0007238443358801305
step: 150, loss: 0.0006540314061567187
step: 160, loss: 9.138359746430069e-05
step: 170, loss: 0.00019394466653466225
step: 180, loss: 0.004147544037550688
step: 190, loss: 5.6047509133350104e-05
step: 200, loss: 3.100755566265434e-05
step: 210, loss: 5.48173047718592e-05
step: 220, loss: 7.147278665797785e-05
step: 230, loss: 0.06047368794679642
/usr/local/lib/python3.7/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=0.9754464285714286, f1=0.9711111111111111,
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: 7.198740058811381e-05
step: 10, loss: 7.531729352194816e-05
step: 20, loss: 7.92374339653179e-05
step: 30, loss: 6.503114127554e-05
step: 40, loss: 0.00011975361121585593
step: 50, loss: 0.00013117125490680337
step: 60, loss: 6.99805241310969e-05
step: 70, loss: 6.694883632007986e-05
step: 80, loss: 8.964839071268216e-05
step: 90, loss: 0.00023276270076166838
step: 100, loss: 6.642071093665436e-05
step: 110, loss: 0.002167933387681842
step: 120, loss: 0.028799254447221756
step: 130, loss: 6.137298623798415e-05
step: 140, loss: 6.870649667689577e-05
step: 150, loss: 9.850790956988931e-05
step: 160, loss: 6.465909245889634e-05
step: 170, loss: 0.0004375089192762971
step: 180, loss: 5.5741562391631305e-05
step: 190, loss: 6.193630542838946e-05
step: 200, loss: 0.0004232917563058436
step: 210, loss: 4.793516200152226e-05
step: 220, loss: 0.001308539416640997
step: 230, loss: 6.19937782175839e-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=0.9754464285714286, f1=0.9711111111111111,
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: 8.665656059747562e-05
step: 10, loss: 0.00016424212662968785
step: 20, loss: 6.415556708816439e-05
step: 30, loss: 0.00011194588296348229
step: 40, loss: 0.012724898755550385
step: 50, loss: 6.124695210019127e-05
step: 60, loss: 8.72412565513514e-05
step: 70, loss: 0.00026866840198636055
step: 80, loss: 5.072389103588648e-05
step: 90, loss: 7.698195986449718e-05
step: 100, loss: 4.877166065853089e-05
step: 110, loss: 8.626198541605845e-05
step: 120, loss: 3.179377017659135e-05
step: 130, loss: 5.144305032445118e-05

```

```

step: 140, loss: 5.362946103559807e-05
step: 150, loss: 4.3255346099613234e-05
step: 160, loss: 0.0006181105272844434
step: 170, loss: 5.381505616242066e-05
step: 180, loss: 0.00012286132550798357
step: 190, loss: 0.0006981941987760365
step: 200, loss: 4.827968223253265e-05
step: 210, loss: 0.002800901886075735
step: 220, loss: 0.00018480255675967783
step: 230, loss: 9.419220441486686e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9754464285714286, f1=0.9743589743589743,
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: 3.756004662136547e-05
step: 10, loss: 3.19110658892896e-05
step: 20, loss: 4.4846205128123984e-05
step: 30, loss: 0.00015690535656176507
step: 40, loss: 0.00011378325871191919
step: 50, loss: 0.052751846611499786
step: 60, loss: 5.196068377699703e-05
step: 70, loss: 6.242353265406564e-05
step: 80, loss: 5.311435234034434e-05
step: 90, loss: 9.762862464413047e-05
step: 100, loss: 7.011849083937705e-05
step: 110, loss: 4.328460636315867e-05
step: 120, loss: 8.638645522296429e-05
step: 130, loss: 0.00037625362165272236
step: 140, loss: 4.691272624768317e-05
step: 150, loss: 0.00012940031592734158
step: 160, loss: 7.653558714082465e-05
step: 170, loss: 0.00017302476044278592
step: 180, loss: 3.153716534143314e-05
step: 190, loss: 5.9757116105174646e-05
step: 200, loss: 4.3427244236227125e-05
step: 210, loss: 0.00010631157056195661
step: 220, loss: 6.603678048122674e-05
step: 230, loss: 6.677297642454505e-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=0.9754464285714286, f1=0.9721913236929923,
best_f1=0.9721913236929923
mv: cannot stat '*_dev.pt': No such file or directory

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

2473it [00:15, 154.82it/s]

load_f1 = 0.9743016759776536

real_f1 = 0.9732739420935412

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

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

4398it [00:25, 174.34it/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: 0.6401538848876953
step: 10, loss: 0.5271133184432983
step: 20, loss: 0.5395477414131165
step: 30, loss: 0.2530089020729065
step: 40, loss: 0.209387868642807
step: 50, loss: 0.10199002921581268
step: 60, loss: 0.12294109165668488
step: 70, loss: 0.10537386685609818
step: 80, loss: 0.07540570199489594
step: 90, loss: 0.26539281010627747
step: 100, loss: 0.06978128105401993
step: 110, loss: 0.0655626729130745
step: 120, loss: 0.11494070291519165
step: 130, loss: 0.10800017416477203
step: 140, loss: 0.07536771148443222
step: 150, loss: 0.021639076992869377
step: 160, loss: 0.1441420167684555
step: 170, loss: 0.27095791697502136
step: 180, loss: 0.06250543892383575
step: 190, loss: 0.02859683893620968
step: 200, loss: 0.08896791189908981
step: 210, loss: 0.09104331582784653
step: 220, loss: 0.32683777809143066
step: 230, loss: 0.2027566283941269
step: 240, loss: 0.07302534580230713
step: 250, loss: 0.03247213363647461
step: 260, loss: 0.043943844735622406
step: 270, loss: 0.0671134740114212
step: 280, loss: 0.094244584441185
step: 290, loss: 0.09308168292045593
step: 300, loss: 0.07471570372581482
step: 310, loss: 0.37389904260635376
step: 320, loss: 0.15756180882453918
step: 330, loss: 0.07670602947473526

```

```

step: 340, loss: 0.06688523292541504
step: 350, loss: 0.007912587374448776
step: 360, loss: 0.08515890687704086
step: 370, loss: 0.12170262634754181
step: 380, loss: 0.07265076786279678
step: 390, loss: 0.23989136517047882
step: 400, loss: 0.2701171636581421
step: 410, loss: 0.1447092443704605
step: 420, loss: 0.03716454282402992
step: 430, loss: 0.12214881181716919
step: 440, loss: 0.019716527312994003
step: 450, loss: 0.015182485803961754
step: 460, loss: 0.002188965678215027
step: 470, loss: 0.18527346849441528
step: 480, loss: 0.20190614461898804
step: 490, loss: 0.0721539705991745
step: 500, loss: 0.3154245615005493
step: 510, loss: 0.06388755887746811
step: 520, loss: 0.07313162088394165
step: 530, loss: 0.002244864124804735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9198473282442748, f1=0.9177365668093201,
best_f1=0.9177365668093201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06201046705245972
step: 10, loss: 0.03548178821802139
step: 20, loss: 0.039283204823732376
step: 30, loss: 0.08089764416217804
step: 40, loss: 0.049668844789266586
step: 50, loss: 0.10941733419895172
step: 60, loss: 0.03163411468267441
step: 70, loss: 0.03564590960741043
step: 80, loss: 0.07304545491933823
step: 90, loss: 0.004586185794323683
step: 100, loss: 0.012803740799427032
step: 110, loss: 0.04239154979586601
step: 120, loss: 0.15766218304634094
step: 130, loss: 0.0419008694589138
step: 140, loss: 0.030303968116641045
step: 150, loss: 0.051222920417785645

```

```

step: 160, loss: 0.02036690153181553
step: 170, loss: 0.1762503832578659
step: 180, loss: 0.011121721006929874
step: 190, loss: 0.026619132608175278
step: 200, loss: 0.04988071694970131
step: 210, loss: 0.0226685032248497
step: 220, loss: 0.04639068990945816
step: 230, loss: 0.041428256779909134
step: 240, loss: 0.027641039341688156
step: 250, loss: 0.030928166583180428
step: 260, loss: 0.0109483003616333
step: 270, loss: 0.1758490949869156
step: 280, loss: 0.13527828454971313
step: 290, loss: 0.05138218402862549
step: 300, loss: 0.19358640909194946
step: 310, loss: 0.004739953204989433
step: 320, loss: 0.17267820239067078
step: 330, loss: 0.034636206924915314
step: 340, loss: 0.01977488212287426
step: 350, loss: 0.005569325760006905
step: 360, loss: 0.0770965963602066
step: 370, loss: 0.0976511761546135
step: 380, loss: 0.06469468772411346
step: 390, loss: 0.05777721107006073
step: 400, loss: 0.05559117719531059
step: 410, loss: 0.005029239691793919
step: 420, loss: 0.01465305034071207
step: 430, loss: 0.01086980476975441
step: 440, loss: 0.14130321145057678
step: 450, loss: 0.019810184836387634
step: 460, loss: 0.06497209519147873
step: 470, loss: 0.0676315501332283
step: 480, loss: 0.18011625111103058
step: 490, loss: 0.04334309324622154
step: 500, loss: 0.1395157277584076
step: 510, loss: 0.010695385746657848
step: 520, loss: 0.04028969258069992
step: 530, loss: 0.008257491514086723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9255813953488372, f1=0.927806241266884,
best_f1=0.927806241266884
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.13477864861488342
step: 10, loss: 0.025313373655080795
step: 20, loss: 0.05544145405292511
step: 30, loss: 0.09590356796979904
step: 40, loss: 0.01427905447781086
step: 50, loss: 0.18011771142482758
step: 60, loss: 0.023379404097795486
step: 70, loss: 0.01605592668056488
step: 80, loss: 0.0036873691715300083
step: 90, loss: 0.008343779481947422
step: 100, loss: 0.06206585094332695
step: 110, loss: 0.010071770288050175
step: 120, loss: 0.0012836792739108205
step: 130, loss: 0.005681023001670837
step: 140, loss: 0.03293886035680771
step: 150, loss: 0.026647862046957016
step: 160, loss: 0.006347024347633123
step: 170, loss: 0.016975073143839836
step: 180, loss: 0.010522155091166496
step: 190, loss: 0.007389269769191742
step: 200, loss: 0.050525516271591187
step: 210, loss: 0.01714610494673252
step: 220, loss: 0.030967634171247482
step: 230, loss: 0.08522655814886093
step: 240, loss: 0.01852487400174141
step: 250, loss: 0.04499583691358566
step: 260, loss: 0.03488089144229889
step: 270, loss: 0.035591065883636475
step: 280, loss: 0.025784309953451157
step: 290, loss: 0.0017195910913869739
step: 300, loss: 0.15764306485652924
step: 310, loss: 0.01570337638258934
step: 320, loss: 0.008696386590600014
step: 330, loss: 0.01095252949744463
step: 340, loss: 0.01075314823538065
step: 350, loss: 0.010321572422981262
step: 360, loss: 0.005225855857133865
step: 370, loss: 0.02231629565358162
step: 380, loss: 0.008234603330492973
step: 390, loss: 0.011051421985030174
step: 400, loss: 0.010936884209513664
step: 410, loss: 0.005891581531614065
step: 420, loss: 0.18603384494781494
step: 430, loss: 0.02645241655409336
step: 440, loss: 0.004203401040285826
step: 450, loss: 0.05596816912293434
```

```
step: 460, loss: 0.021073097363114357
step: 470, loss: 0.12960900366306305
step: 480, loss: 0.003390482859686017
step: 490, loss: 0.0026894290931522846
step: 500, loss: 0.0026594032533466816
step: 510, loss: 0.004874768666923046
step: 520, loss: 0.02077385038137436
step: 530, loss: 0.1318499594926834
/usr/local/lib/python3.7/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=0.9217230199166282, f1=0.9245810055865922,
best_f1=0.927806241266884
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045502983033657074
step: 10, loss: 0.011053884401917458
step: 20, loss: 0.002392670139670372
step: 30, loss: 0.0020689694210886955
step: 40, loss: 0.011273362673819065
step: 50, loss: 0.008522656746208668
step: 60, loss: 0.0013869507238268852
step: 70, loss: 0.01089287456125021
step: 80, loss: 0.0034949248656630516
step: 90, loss: 0.05511876568198204
step: 100, loss: 0.053844012320041656
step: 110, loss: 0.0009906886843964458
step: 120, loss: 0.0008594093378633261
step: 130, loss: 0.0021289505530148745
step: 140, loss: 0.0021750726737082005
step: 150, loss: 0.0022405986674129963
step: 160, loss: 0.02304132841527462
step: 170, loss: 0.009213045239448547
step: 180, loss: 0.0008210256928578019
step: 190, loss: 0.009083117358386517
step: 200, loss: 0.004377780016511679
step: 210, loss: 0.03085250034928322
step: 220, loss: 0.07536693662405014
step: 230, loss: 0.046315521001815796
step: 240, loss: 0.010189064778387547
step: 250, loss: 0.03920721635222435
step: 260, loss: 0.007785438559949398
step: 270, loss: 0.016062874346971512
step: 280, loss: 0.052552834153175354
step: 290, loss: 0.012539029121398926
step: 300, loss: 0.002102301688864827
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```

step: 310, loss: 0.09568502008914948
step: 320, loss: 0.013332941569387913
step: 330, loss: 0.014418313279747963
step: 340, loss: 0.012880860827863216
step: 350, loss: 0.004607740323990583
step: 360, loss: 0.09032897651195526
step: 370, loss: 0.002905243309214711
step: 380, loss: 0.005728668533265591
step: 390, loss: 0.01897045411169529
step: 400, loss: 0.02726600132882595
step: 410, loss: 0.003148965537548065
step: 420, loss: 0.0025248529855161905
step: 430, loss: 0.23899289965629578
step: 440, loss: 0.025935662910342216
step: 450, loss: 0.005943263880908489
step: 460, loss: 0.001816740958020091
step: 470, loss: 0.008754117414355278
step: 480, loss: 0.07322521507740021
step: 490, loss: 0.0033673145808279514
step: 500, loss: 0.006238862406462431
step: 510, loss: 0.0038687465712428093
step: 520, loss: 0.034670691937208176
step: 530, loss: 0.04705338180065155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9268965517241379, f1=0.9191176470588236,
best_f1=0.9191176470588236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022188616916537285
step: 10, loss: 0.05883782356977463
step: 20, loss: 0.003287564031779766
step: 30, loss: 0.001476478879339993
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.0035209658090025187
step: 50, loss: 0.002397894859313965
step: 60, loss: 0.1042519062757492
step: 70, loss: 0.015054145827889442
step: 80, loss: 0.000991565641015768
step: 90, loss: 0.12588514387607574
step: 100, loss: 0.002615648088976741
step: 110, loss: 0.007512697484344244

```

```
step: 120, loss: 0.00328287435695529
step: 130, loss: 0.002181215211749077
step: 140, loss: 0.012645849958062172
step: 150, loss: 0.000684526632539928
step: 160, loss: 0.0006714833434671164
step: 170, loss: 0.0036154098343104124
step: 180, loss: 0.0019811862148344517
step: 190, loss: 0.04252588376402855
step: 200, loss: 0.005292113870382309
step: 210, loss: 0.05448869243264198
step: 220, loss: 0.0005627591162919998
step: 230, loss: 0.0029295969288796186
step: 240, loss: 0.0008681846666149795
step: 250, loss: 0.0009456724510528147
step: 260, loss: 0.007664625067263842
step: 270, loss: 0.00042560204747132957
step: 280, loss: 0.013500327244400978
step: 290, loss: 0.12938112020492554
step: 300, loss: 0.000837442115880549
step: 310, loss: 0.004045713227242231
step: 320, loss: 0.08994980901479721
step: 330, loss: 0.008183175697922707
step: 340, loss: 0.0070104338228702545
step: 350, loss: 0.0015418839175254107
step: 360, loss: 0.004561872687190771
step: 370, loss: 0.00919525045901537
step: 380, loss: 0.00746507104486227
step: 390, loss: 0.0019615362398326397
step: 400, loss: 0.004757968243211508
step: 410, loss: 0.05512811988592148
step: 420, loss: 0.002488648286089301
step: 430, loss: 0.011598451994359493
step: 440, loss: 0.0009751128382049501
step: 450, loss: 0.01512247696518898
step: 460, loss: 0.06826120615005493
step: 470, loss: 0.00021692723385058343
step: 480, loss: 0.005065407603979111
step: 490, loss: 0.0010537991765886545
step: 500, loss: 0.01423073373734951
step: 510, loss: 0.10451523959636688
step: 520, loss: 0.003289534943178296
step: 530, loss: 0.002538889180868864
/usr/local/lib/python3.7/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=0.9190225203641591, f1=0.9111969111969113,
best_f1=0.9191176470588236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```


An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006449423381127417
step: 10, loss: 0.0003680954105220735
step: 20, loss: 0.0589577816426754
step: 30, loss: 0.019729170948266983
step: 40, loss: 0.002875187899917364
step: 50, loss: 0.015264442190527916
step: 60, loss: 0.0009090523817576468
step: 70, loss: 0.002078230492770672
step: 80, loss: 0.022494060918688774
step: 90, loss: 0.0029107369482517242
step: 100, loss: 0.005333289969712496
step: 110, loss: 0.013210305012762547
step: 120, loss: 0.004757918417453766
step: 130, loss: 0.028911909088492393
step: 140, loss: 0.03035266324877739
step: 150, loss: 0.0003103116760030389
step: 160, loss: 0.0003648687561508268
step: 170, loss: 0.004604641813784838
step: 180, loss: 0.0002901928673963994
step: 190, loss: 0.002520928857848048
step: 200, loss: 0.0005120753776282072
step: 210, loss: 0.002474858658388257
step: 220, loss: 0.0014927223091945052
step: 230, loss: 0.00027030243654735386
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 240, loss: 0.02985476702451706
step: 250, loss: 0.007041182368993759
step: 260, loss: 0.0002211560058640316
step: 270, loss: 0.03962944447994232
step: 280, loss: 0.005947644356638193
step: 290, loss: 0.0007690790807828307
step: 300, loss: 0.00032288339571096003
step: 310, loss: 0.0015869186026975513
step: 320, loss: 0.00046870720689184964
step: 330, loss: 0.00031643389957025647
step: 340, loss: 0.1848650723695755
step: 350, loss: 0.002311017131432891
step: 360, loss: 0.003658146597445011
step: 370, loss: 0.005669057834893465
step: 380, loss: 0.024263115599751472
step: 390, loss: 0.04082760587334633
step: 400, loss: 0.0005967046599835157
step: 410, loss: 0.0016596560599282384
step: 420, loss: 0.003647084115073085
step: 430, loss: 0.001041869167238474
step: 440, loss: 9.09975788090378e-05
```

```

step: 450, loss: 0.00028291920898482203
step: 460, loss: 0.00023708517255727202
step: 470, loss: 0.0004952634335495532
step: 480, loss: 0.001037185313180089
step: 490, loss: 0.00697718420997262
step: 500, loss: 0.0017517883097752929
step: 510, loss: 0.00316020124591887
step: 520, loss: 0.00038662305450998247
step: 530, loss: 0.004147205501794815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.929472209248015, f1=0.922429906542056,
best_f1=0.922429906542056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002953251823782921
step: 10, loss: 0.0010351671371608973
step: 20, loss: 0.00023492488253396004
step: 30, loss: 0.000992572051472962
step: 40, loss: 0.0022657897789031267
step: 50, loss: 0.0004619878891389817
step: 60, loss: 0.00025149897555820644
step: 70, loss: 0.0002734569425228983
step: 80, loss: 0.0011926733423024416
step: 90, loss: 0.00011846820416394621
step: 100, loss: 0.0001382593618473038
step: 110, loss: 0.0005337856709957123
step: 120, loss: 0.00045462572597898543
step: 130, loss: 0.001461649313569069
step: 140, loss: 0.0001598942617420107
step: 150, loss: 0.00028210534946992993
step: 160, loss: 0.00010849224781850353
step: 170, loss: 0.0001938873465405777
step: 180, loss: 0.0031760758720338345
step: 190, loss: 0.001503786537796259
step: 200, loss: 0.0007829376263543963
step: 210, loss: 0.0062194219790399075
step: 220, loss: 0.00021523858595173806
step: 230, loss: 0.0011152623919770122
step: 240, loss: 0.06398449093103409
step: 250, loss: 0.0020279441960155964
step: 260, loss: 0.00021343218395486474

```

```

step: 270, loss: 0.011989885941147804
step: 280, loss: 0.0017484648851677775
step: 290, loss: 0.0002052186755463481
step: 300, loss: 0.0021055140532553196
step: 310, loss: 0.0002376854681642726
step: 320, loss: 0.00011328660184517503
step: 330, loss: 0.00015824708680156618
step: 340, loss: 0.013145903125405312
step: 350, loss: 0.002998793264850974
step: 360, loss: 0.0005743541987612844
step: 370, loss: 0.0002519587869755924
step: 380, loss: 0.0006631025462411344
step: 390, loss: 0.0005482654669322073
step: 400, loss: 0.001682364963926375
step: 410, loss: 0.0029505181591957808
step: 420, loss: 0.08443023264408112
step: 430, loss: 0.000259123626165092
step: 440, loss: 0.0011031870963051915
step: 450, loss: 0.0013710687635466456
step: 460, loss: 0.0011168767232447863
step: 470, loss: 0.031873028725385666
step: 480, loss: 0.2735539376735687
step: 490, loss: 0.00311130378395319
step: 500, loss: 0.004151828121393919
step: 510, loss: 0.0010296439286321402
step: 520, loss: 0.06744657456874847
step: 530, loss: 0.01127712707966566
/usr/local/lib/python3.7/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=0.9255514705882354, f1=0.9249884951679707,
best_f1=0.922429906542056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017118489369750023
step: 10, loss: 0.029863614588975906
step: 20, loss: 0.01031381543725729
step: 30, loss: 0.004430266097187996
step: 40, loss: 0.00019953883020207286
step: 50, loss: 0.007979612797498703
step: 60, loss: 0.0024538501165807247
step: 70, loss: 0.00026062096003443
step: 80, loss: 0.001624319120310247
step: 90, loss: 0.002265307819470763
step: 100, loss: 0.0001698596024652943
step: 110, loss: 0.0001601121766725555
step: 120, loss: 0.0010157402139157057

```

```

step: 130, loss: 5.503366264747456e-05
step: 140, loss: 0.0007515526376664639
step: 150, loss: 8.86302295839414e-05
step: 160, loss: 6.897825369378552e-05
step: 170, loss: 0.0022159155923873186
step: 180, loss: 7.951632142066956e-05
step: 190, loss: 0.003278203774243593
step: 200, loss: 0.001459159073419869
step: 210, loss: 0.0004446603707037866
step: 220, loss: 0.0015708478167653084
step: 230, loss: 0.001870832871645689
step: 240, loss: 6.597326137125492e-05
step: 250, loss: 9.98240357148461e-05
step: 260, loss: 0.00045359483920037746
step: 270, loss: 0.00025999703211709857
step: 280, loss: 0.021626897156238556
step: 290, loss: 0.02724933996796608
step: 300, loss: 0.000561566324904561
step: 310, loss: 0.000547628675121814
step: 320, loss: 0.00731453113257885
step: 330, loss: 0.004877671133726835
step: 340, loss: 0.004052099771797657
step: 350, loss: 0.00016785661864560097
step: 360, loss: 0.005841994658112526
step: 370, loss: 0.010775534436106682
step: 380, loss: 0.0007940199575386941
step: 390, loss: 0.08088970929384232
step: 400, loss: 0.0009426175965927541
step: 410, loss: 0.018300721421837807
step: 420, loss: 0.19054046273231506
step: 430, loss: 0.01688367687165737
step: 440, loss: 0.0013660952681675553
step: 450, loss: 0.002607084810733795
step: 460, loss: 0.00014144399028737098
step: 470, loss: 0.002184620127081871
step: 480, loss: 0.0001966389245353639
step: 490, loss: 0.0012822265271097422
step: 500, loss: 0.007905692793428898
step: 510, loss: 0.0001091706653824076
step: 520, loss: 0.002888580085709691
step: 530, loss: 0.00058348587481305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9298653042266605, f1=0.9285051067780873,  
best_f1=0.9285051067780873  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0010689833434298635  
step: 10, loss: 8.95058547030203e-05  
step: 20, loss: 0.0004760146839544177  
step: 30, loss: 0.00018343847477808595  
step: 40, loss: 6.974796269787475e-05  
step: 50, loss: 9.458001295570284e-05  
step: 60, loss: 0.00010373591067036614  
step: 70, loss: 0.0032674826215952635  
step: 80, loss: 0.00016368307115044445  
step: 90, loss: 0.0001856388698797673  
step: 100, loss: 0.00022157050261739641  
step: 110, loss: 4.790012098965235e-05  
step: 120, loss: 0.002873822348192334  
step: 130, loss: 8.015539788175374e-05  
step: 140, loss: 0.0009776361985132098  
step: 150, loss: 0.0001247016916749999  
step: 160, loss: 0.00046584507799707353  
step: 170, loss: 0.001523195649497211  
step: 180, loss: 3.912073952960782e-05  
step: 190, loss: 0.0001949152210727334  
step: 200, loss: 0.00022426896612159908  
step: 210, loss: 0.002677571028470993  
step: 220, loss: 0.00014753038703929633  
step: 230, loss: 6.568247772520408e-05  
step: 240, loss: 0.00011494516365928575  
step: 250, loss: 0.001879124785773456  
step: 260, loss: 0.00029648770578205585  
step: 270, loss: 3.89717570214998e-05  
step: 280, loss: 4.268877091817558e-05  
step: 290, loss: 0.12357119470834732  
step: 300, loss: 7.252550858538598e-05  
step: 310, loss: 0.0006145605002529919  
step: 320, loss: 0.0002383867249591276  
step: 330, loss: 0.002311249729245901  
step: 340, loss: 0.0015645261155441403  
step: 350, loss: 0.00015915681433398277  
step: 360, loss: 0.00019273637735750526  
step: 370, loss: 7.587856089230627e-05  
step: 380, loss: 0.0023818337358534336  
step: 390, loss: 7.822853513062e-05  
step: 400, loss: 0.001662084017880261  
step: 410, loss: 6.153889262350276e-05  
step: 420, loss: 6.402089638868347e-05
```

```

step: 430, loss: 0.00016927049728110433
step: 440, loss: 0.0005607751663774252
step: 450, loss: 8.972528303274885e-05
step: 460, loss: 0.000203123883693479
step: 470, loss: 3.2695665140636265e-05
step: 480, loss: 5.440023960545659e-05
step: 490, loss: 0.007807765156030655
step: 500, loss: 0.00048523908481001854
step: 510, loss: 0.00010282552102580667
step: 520, loss: 7.220383122330531e-05
step: 530, loss: 0.00010145410487893969
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9316081330868762, f1=0.9225746268656716,
best_f1=0.9225746268656716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001199384787469171
step: 10, loss: 9.716282511362806e-05
step: 20, loss: 3.408172415220179e-05
step: 30, loss: 4.207864549243823e-05
step: 40, loss: 6.263165414566174e-05
step: 50, loss: 0.010829826816916466
step: 60, loss: 7.134562474675477e-05
step: 70, loss: 4.837234519072808e-05
step: 80, loss: 7.480860222131014e-05
step: 90, loss: 7.586805440951139e-05
step: 100, loss: 0.0003710029413923621
step: 110, loss: 4.690060450229794e-05
step: 120, loss: 4.6887937060091645e-05
step: 130, loss: 7.162632391555235e-05
step: 140, loss: 0.0005728536634705961
step: 150, loss: 3.950099562644027e-05
step: 160, loss: 3.7980396882630885e-05
step: 170, loss: 0.00020582045544870198
step: 180, loss: 6.362967542372644e-05
step: 190, loss: 6.185924576129764e-05
step: 200, loss: 0.00010862851922865957
step: 210, loss: 0.0002811908780131489
step: 220, loss: 6.992560520302504e-05
step: 230, loss: 6.581060006283224e-05
step: 240, loss: 0.04364767298102379

```

```

step: 250, loss: 8.477405935991555e-05
step: 260, loss: 0.0019024558132514358
step: 270, loss: 0.001685073715634644
step: 280, loss: 9.457394480705261e-05
step: 290, loss: 0.0001282378943869844
step: 300, loss: 6.166086677694693e-05
step: 310, loss: 9.320351819042116e-05
step: 320, loss: 0.0014931083424016833
step: 330, loss: 3.704212576849386e-05
step: 340, loss: 0.00015080076991580427
step: 350, loss: 0.0005966787575744092
step: 360, loss: 9.356747614219785e-05
step: 370, loss: 0.0009510516538284719
step: 380, loss: 0.003025380428880453
step: 390, loss: 9.51049369177781e-05
step: 400, loss: 0.00146497692912817
step: 410, loss: 0.0008401491795666516
step: 420, loss: 0.0002553588419687003
step: 430, loss: 5.8082208852283657e-05
step: 440, loss: 0.020656144246459007
step: 450, loss: 0.0007170903263613582
step: 460, loss: 5.984727249597199e-05
step: 470, loss: 0.0001475926983403042
step: 480, loss: 7.059828203637153e-05
step: 490, loss: 0.0017948005115613341
step: 500, loss: 0.0015228809788823128
step: 510, loss: 0.00021685686078853905
step: 520, loss: 0.0002492867934051901
step: 530, loss: 0.0015314146876335144
/usr/local/lib/python3.7/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=0.9291044776119404, f1=0.9198508853681266,
best_f1=0.9225746268656716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001442168722860515
step: 10, loss: 0.00029975443612784147
step: 20, loss: 0.0009200314525514841
step: 30, loss: 0.00034478199086152017
step: 40, loss: 0.0001549317967146635
step: 50, loss: 0.0009318897500634193
step: 60, loss: 0.0016508768312633038
step: 70, loss: 9.218810009770095e-05
step: 80, loss: 0.00013750455400440842
step: 90, loss: 0.00018207027460448444
step: 100, loss: 0.0008513516513630748

```

```
step: 110, loss: 0.0016665119910612702
step: 120, loss: 4.3980780901620165e-05
step: 130, loss: 6.694552575936541e-05
step: 140, loss: 0.000683921913150698
step: 150, loss: 0.00020310045510996133
step: 160, loss: 5.597972631221637e-05
step: 170, loss: 8.920658729039133e-05
step: 180, loss: 0.0001729264622554183
step: 190, loss: 0.00014629164070356637
step: 200, loss: 0.0001080253641703166
step: 210, loss: 0.00010661972919479012
step: 220, loss: 0.0004290277720429003
step: 230, loss: 3.6823224945692345e-05
step: 240, loss: 0.00013551664596889168
step: 250, loss: 4.154441921855323e-05
step: 260, loss: 4.733721289085224e-05
step: 270, loss: 5.715616498491727e-05
step: 280, loss: 8.366024849237874e-05
step: 290, loss: 0.0005357501213438809
step: 300, loss: 7.436168380081654e-05
step: 310, loss: 0.0008205648045986891
step: 320, loss: 5.317037721397355e-05
step: 330, loss: 0.0001916781038744375
step: 340, loss: 0.0004030222480650991
step: 350, loss: 0.0017156670801341534
step: 360, loss: 5.217011857894249e-05
step: 370, loss: 0.01712566427886486
step: 380, loss: 0.0004614498757291585
step: 390, loss: 0.0022526136599481106
step: 400, loss: 0.0007556856144219637
step: 410, loss: 0.0001785576605470851
step: 420, loss: 0.002385316649451852
step: 430, loss: 0.0009538480662740767
step: 440, loss: 0.0192805677652359
step: 450, loss: 5.9182904806220904e-05
step: 460, loss: 0.00781718548387289
step: 470, loss: 0.0003148595569655299
step: 480, loss: 0.0003437626291997731
step: 490, loss: 7.853434362914413e-05
step: 500, loss: 0.0003470649535302073
step: 510, loss: 5.2673978643724695e-05
step: 520, loss: 0.00020265727653168142
step: 530, loss: 0.00014370103599503636
/usr/local/lib/python3.7/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=0.9263360616273472, f1=0.9155812831644958,
best_f1=0.9225746268656716
```



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

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

```
step: 0, loss: 0.0001486387918703258  
step: 10, loss: 3.9973907405510545e-05  
step: 20, loss: 7.446197560057044e-05  
step: 30, loss: 0.00011496577644720674  
step: 40, loss: 0.00013832775584887713  
step: 50, loss: 0.02108158729970455  
step: 60, loss: 0.003274462418630719  
step: 70, loss: 0.00013844580098520964  
step: 80, loss: 0.00024969433434307575  
step: 90, loss: 8.585509203840047e-05  
step: 100, loss: 5.2847452025162056e-05  
step: 110, loss: 0.00013165888958610594  
step: 120, loss: 0.00016157046775333583  
step: 130, loss: 6.052304161130451e-05  
step: 140, loss: 0.0012778742238879204  
step: 150, loss: 2.976998257508967e-05  
step: 160, loss: 0.00033101870212703943  
step: 170, loss: 9.185768431052566e-05  
step: 180, loss: 0.00014103377179708332  
step: 190, loss: 5.097346365801059e-05  
step: 200, loss: 8.879615052137524e-05  
step: 210, loss: 0.00010315689723938704  
step: 220, loss: 0.054803598672151566  
step: 230, loss: 6.116056465543807e-05  
step: 240, loss: 0.04043520614504814  
step: 250, loss: 0.00011788264237111434  
step: 260, loss: 4.691235153586604e-05  
step: 270, loss: 8.760247874306515e-05  
step: 280, loss: 0.07718755304813385  
step: 290, loss: 5.070135375717655e-05  
step: 300, loss: 0.0001178421953227371  
step: 310, loss: 0.0002826599811669439  
step: 320, loss: 2.8634896807488985e-05  
step: 330, loss: 0.0001201246413984336  
step: 340, loss: 3.805574669968337e-05  
step: 350, loss: 0.00011545919551281258  
step: 360, loss: 0.0298676285892725  
step: 370, loss: 0.0001131410535890609  
step: 380, loss: 0.0001526612468296662  
step: 390, loss: 0.00028613951872102916  
step: 400, loss: 0.00021017494145780802  
step: 410, loss: 0.00014596444088965654  
step: 420, loss: 0.005347154103219509  
step: 430, loss: 0.00012828284525312483  
step: 440, loss: 0.01374081801623106
```

```
step: 450, loss: 0.0013035460142418742
step: 460, loss: 0.00018974480917677283
step: 470, loss: 6.970611138967797e-05
step: 480, loss: 0.000803899485617876
step: 490, loss: 0.001172534073702991
step: 500, loss: 6.806878809584305e-05
step: 510, loss: 0.00016702592256478965
step: 520, loss: 6.533100531669334e-05
step: 530, loss: 0.00010621146793710068
/usr/local/lib/python3.7/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=0.9278642149929278, f1=0.9200945626477541,
best_f1=0.9225746268656716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019202566472813487
step: 10, loss: 9.150293772108853e-05
step: 20, loss: 2.2153564714244567e-05
step: 30, loss: 0.00010873060818994418
step: 40, loss: 3.0103001336101443e-05
step: 50, loss: 0.003647561650723219
step: 60, loss: 3.551463669282384e-05
step: 70, loss: 7.47641606722027e-05
step: 80, loss: 0.00017598233534954488
step: 90, loss: 0.00041824017534963787
step: 100, loss: 4.852686106460169e-05
step: 110, loss: 2.910364310082514e-05
step: 120, loss: 3.0091412554611452e-05
step: 130, loss: 2.4094375476124696e-05
step: 140, loss: 8.09090633993037e-05
step: 150, loss: 0.0008061896078288555
step: 160, loss: 4.3163716327399015e-05
step: 170, loss: 4.278985579730943e-05
step: 180, loss: 0.0002782979281619191
step: 190, loss: 7.269695197464898e-05
step: 200, loss: 0.00018879343406297266
step: 210, loss: 0.00015729782171547413
step: 220, loss: 0.00011187530617462471
step: 230, loss: 3.252373790019192e-05
step: 240, loss: 4.426849773153663e-05
step: 250, loss: 4.1117080399999395e-05
step: 260, loss: 4.1392268030904233e-05
step: 270, loss: 2.1274438040563837e-05
step: 280, loss: 7.378587179118767e-05
step: 290, loss: 0.003264690050855279
step: 300, loss: 0.00014635425759479403
```

```

step: 310, loss: 0.0005545567255467176
step: 320, loss: 0.0023326745722442865
step: 330, loss: 0.0011963449651375413
step: 340, loss: 9.874007082544267e-05
step: 350, loss: 0.0003455231199041009
step: 360, loss: 2.647102519404143e-05
step: 370, loss: 9.116619185078889e-05
step: 380, loss: 0.00010041419591289014
step: 390, loss: 0.04853369668126106
step: 400, loss: 0.0009842927101999521
step: 410, loss: 0.0001336423447355628
step: 420, loss: 4.585032365866937e-05
step: 430, loss: 0.0006814097869209945
step: 440, loss: 3.927223224309273e-05
step: 450, loss: 5.480658728629351e-05
step: 460, loss: 0.0031743193976581097
step: 470, loss: 0.0005329891573637724
step: 480, loss: 3.640287832240574e-05
step: 490, loss: 3.198771082679741e-05
step: 500, loss: 3.797331373789348e-05
step: 510, loss: 0.00033563192118890584
step: 520, loss: 5.110584243084304e-05
step: 530, loss: 0.00025098599144257605
/usr/local/lib/python3.7/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=0.92524682651622, f1=0.9222846441947565,
best_f1=0.9225746268656716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000848327879793942
step: 10, loss: 2.3762320779496804e-05
step: 20, loss: 1.9806962882285006e-05
step: 30, loss: 3.5384651710046455e-05
step: 40, loss: 5.971762948320247e-05
step: 50, loss: 0.00038319238228723407
step: 60, loss: 2.39415057876613e-05
step: 70, loss: 3.609043778851628e-05
step: 80, loss: 3.9480975829064846e-05
step: 90, loss: 7.587313302792609e-05
step: 100, loss: 3.651336737675592e-05
step: 110, loss: 2.6802617867360823e-05
step: 120, loss: 2.295094782311935e-05
step: 130, loss: 0.004856090992689133
step: 140, loss: 0.00025383010506629944
step: 150, loss: 2.964660234283656e-05
step: 160, loss: 2.687336200324353e-05

```

```

step: 170, loss: 4.2930398194584996e-05
step: 180, loss: 3.188647315255366e-05
step: 190, loss: 3.430020296946168e-05
step: 200, loss: 7.343583274632692e-05
step: 210, loss: 3.833406663034111e-05
step: 220, loss: 2.9875400286982767e-05
step: 230, loss: 0.0014031408354640007
step: 240, loss: 3.246391861466691e-05
step: 250, loss: 3.6327357520349324e-05
step: 260, loss: 5.125040843267925e-05
step: 270, loss: 8.364295354112983e-05
step: 280, loss: 2.6623640223988332e-05
step: 290, loss: 0.006346886977553368
step: 300, loss: 4.700299177784473e-05
step: 310, loss: 0.0014110878109931946
step: 320, loss: 0.0001892985455924645
step: 330, loss: 0.00012807831808459014
step: 340, loss: 0.001196966040879488
step: 350, loss: 0.00015871234063524753
step: 360, loss: 0.0274994857609272
step: 370, loss: 2.5115205062320456e-05
step: 380, loss: 6.397102697519585e-05
step: 390, loss: 0.003624941688030958
step: 400, loss: 0.00016193697229027748
step: 410, loss: 2.1043448214186355e-05
step: 420, loss: 4.425036604516208e-05
step: 430, loss: 0.00021534493134822696
step: 440, loss: 5.045751458965242e-05
step: 450, loss: 4.5221222535474226e-05
step: 460, loss: 0.00011330330016789958
step: 470, loss: 4.550943413050845e-05
step: 480, loss: 2.9804790756315924e-05
step: 490, loss: 0.007428786251693964
step: 500, loss: 3.388606637599878e-05
step: 510, loss: 0.0001841234916355461
step: 520, loss: 8.421538950642571e-05
step: 530, loss: 3.4485205105738714e-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=0.930562116202173, f1=0.9210526315789475,
best_f1=0.9225746268656716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6526706278673373e-05
step: 10, loss: 4.554905171971768e-05
step: 20, loss: 1.7337208191747777e-05

```

step: 30, loss: 5.265435902401805e-05
step: 40, loss: 0.01608952507376671
step: 50, loss: 6.236953049665317e-05
step: 60, loss: 1.6349951692973264e-05
step: 70, loss: 0.0002144872851204127
step: 80, loss: 0.0020097438246011734
step: 90, loss: 2.4992303224280477e-05
step: 100, loss: 0.00016214251809287816
step: 110, loss: 0.0002704719372559339
step: 120, loss: 8.375842298846692e-05
step: 130, loss: 0.046496883034706116
step: 140, loss: 3.0499024433083832e-05
step: 150, loss: 6.519411545014009e-05
step: 160, loss: 4.954488031216897e-05
step: 170, loss: 4.8988535127136856e-05
step: 180, loss: 1.617490670469124e-05
step: 190, loss: 7.090610597515479e-05
step: 200, loss: 2.1277914129314013e-05
step: 210, loss: 2.3516633518738672e-05
step: 220, loss: 2.9227245249785483e-05
step: 230, loss: 2.752531690930482e-05
step: 240, loss: 3.152454155497253e-05
step: 250, loss: 3.200228456989862e-05
step: 260, loss: 2.8735723390127532e-05
step: 270, loss: 2.5472902052570134e-05
step: 280, loss: 2.9797532988595776e-05
step: 290, loss: 6.165917875478044e-05
step: 300, loss: 2.616490382933989e-05
step: 310, loss: 0.0016024911310523748
step: 320, loss: 1.7556971215526573e-05
step: 330, loss: 4.0510469261789694e-05
step: 340, loss: 1.7258944353670813e-05
step: 350, loss: 1.5858311599004082e-05
step: 360, loss: 0.00020701820903923362
step: 370, loss: 3.605450547183864e-05
step: 380, loss: 5.4972759244265035e-05
step: 390, loss: 0.00018906490004155785
step: 400, loss: 2.5506302335998043e-05
step: 410, loss: 1.724410140013788e-05
step: 420, loss: 7.706562610110268e-05
step: 430, loss: 4.0890892705647275e-05
step: 440, loss: 0.0001333348627667874
step: 450, loss: 2.0767894966411404e-05
step: 460, loss: 3.2225514587480575e-05
step: 470, loss: 0.00015869484923314303
step: 480, loss: 1.6323872841894627e-05
step: 490, loss: 1.9609506125561893e-05
step: 500, loss: 7.066094258334488e-05

```

step: 510, loss: 2.6004847313743085e-05
step: 520, loss: 2.995379691128619e-05
step: 530, loss: 3.429602293181233e-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=0.9286047596826879, f1=0.9208566108007449,
best_f1=0.9225746268656716
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 [00:32, 176.53it/s]
load_f1 = 0.9304713019132057
real_f1 = 0.9259431765253843
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 177.78it/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: 0.5642581582069397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.36, f1=0.31250000000000006, best_f1=0.31250000000000006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48715269565582275
/usr/local/lib/python3.7/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=0.33802816901408456, f1=0.32, best_f1=0.31250000000000006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5358408093452454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4285714285714286, f1=0.35714285714285715,
best_f1=0.35714285714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26885199546813965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.55, f1=0.4666666666666667, best_f1=0.4666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2420109212398529
/usr/local/lib/python3.7/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=0.5238095238095237, f1=0.5000000000000001,
best_f1=0.4666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.290209025144577
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6250000000000001, f1=0.6000000000000001,
best_f1=0.6000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.07456933706998825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6470588235294117, f1=0.5625000000000001,
best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027730198577046394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7096774193548386, f1=0.5384615384615384,
best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036831044126302004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.6666666666666667, f1=0.5517241379310344,
best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01122783962637186
/usr/local/lib/python3.7/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=0.7096774193548386, f1=0.6428571428571429,
best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003516497788950801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7333333333333334, 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: 0.008601327426731586
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.6666666666666666,
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: 0.003398618195205927
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.6875000000000001,
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: 0.040497712790966034
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.6875000000000001,
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: 0.005001570098102093
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.6875000000000001,
best_f1=0.7142857142857143
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, 130623.43it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.5714285714285714
real_f1 = 0.5517241379310344
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 179.09it/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/

```

```

Downloading: 100% 433/433 [00:00<00:00, 259kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 3.05MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 64.1MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
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: 0.6309524774551392
step: 10, loss: 0.6353327631950378
step: 20, loss: 0.36955833435058594
step: 30, loss: 0.13447456061840057
step: 40, loss: 0.3101920485496521
step: 50, loss: 0.02532215043902397
step: 60, loss: 0.08127039670944214
step: 70, loss: 0.07467811554670334
step: 80, loss: 0.16405801475048065
step: 90, loss: 0.12084567546844482
step: 100, loss: 0.007158154621720314
step: 110, loss: 0.1961156725883484
step: 120, loss: 0.0049243527464568615
step: 130, loss: 0.029751237481832504
step: 140, loss: 0.005904721561819315
step: 150, loss: 0.079658642411232
step: 160, loss: 0.024897560477256775
step: 170, loss: 0.09821554273366928
step: 180, loss: 0.014255683869123459
step: 190, loss: 0.012995891273021698
step: 200, loss: 0.017972638830542564
step: 210, loss: 0.01534099131822586
step: 220, loss: 0.015442471951246262
step: 230, loss: 0.10972655564546585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9702517162471395, f1=0.9714937286202965,
best_f1=0.9714937286202965
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003927348647266626
step: 10, loss: 0.004230109043419361
step: 20, loss: 0.08434463292360306
step: 30, loss: 0.20313726365566254
step: 40, loss: 0.035366058349609375
step: 50, loss: 0.02924496866762638
step: 60, loss: 0.03061888925731182
step: 70, loss: 0.2220505326986313
step: 80, loss: 0.004465845879167318
step: 90, loss: 0.004663174040615559
step: 100, loss: 0.04422343149781227
step: 110, loss: 0.11124139279127121
step: 120, loss: 0.01538803894072771
step: 130, loss: 0.004771477077156305
step: 140, loss: 0.024769747629761696
step: 150, loss: 0.025295615196228027
step: 160, loss: 0.019113842397928238
step: 170, loss: 0.0038158465176820755
step: 180, loss: 0.0032693215180188417
step: 190, loss: 0.04636864364147186
step: 200, loss: 0.009639403782784939
step: 210, loss: 0.0007935917819850147
step: 220, loss: 0.2882872223854065
step: 230, loss: 0.025356708094477654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.9819819819819819,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00803819578140974
step: 10, loss: 0.0013860843610018492
step: 20, loss: 0.0008520131814293563
step: 30, loss: 0.1120707169175148
step: 40, loss: 0.1366945505142212
step: 50, loss: 0.013126278296113014
step: 60, loss: 0.01499294675886631
step: 70, loss: 0.0026436520274728537
step: 80, loss: 0.002001016866415739
step: 90, loss: 0.07899680733680725

```

```

step: 100, loss: 0.0013641545083373785
step: 110, loss: 0.0006097399746067822
step: 120, loss: 0.0167809147387743
step: 130, loss: 0.007214122451841831
step: 140, loss: 0.018978416919708252
step: 150, loss: 0.022443639114499092
step: 160, loss: 0.0578186996281147
step: 170, loss: 0.004396816249936819
step: 180, loss: 0.0035577628295868635
step: 190, loss: 0.0028757695108652115
step: 200, loss: 0.05102790892124176
step: 210, loss: 0.0008730792324058712
step: 220, loss: 0.0008424895931966603
step: 230, loss: 0.0004099791403859854
/usr/local/lib/python3.7/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=0.9865168539325843, f1=0.977728285077951,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014279666356742382
step: 10, loss: 0.0003278766816947609
step: 20, loss: 0.0007981530507095158
step: 30, loss: 0.004593153949826956
step: 40, loss: 0.005910373758524656
step: 50, loss: 0.005807633511722088
step: 60, loss: 0.0006409790948964655
step: 70, loss: 0.0005715576116926968
step: 80, loss: 0.0004868078394792974
step: 90, loss: 0.0005929536418989301
step: 100, loss: 0.0007845099898986518
step: 110, loss: 0.0003771836927626282
step: 120, loss: 0.15854839980602264
step: 130, loss: 0.0082571255043149
step: 140, loss: 0.0009944919729605317
step: 150, loss: 0.16535557806491852
step: 160, loss: 0.0006451947847381234
step: 170, loss: 0.01774054393172264
step: 180, loss: 0.0004714437818620354
step: 190, loss: 0.009331990033388138
step: 200, loss: 0.0026158036198467016
step: 210, loss: 0.010381920263171196
step: 220, loss: 0.0005626103957183659
step: 230, loss: 0.01326585840433836
/usr/local/lib/python3.7/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=0.9775280898876404, f1=0.9764309764309763,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027336194179952145
step: 10, loss: 0.0004430878034327179
step: 20, loss: 0.000786366465035826
step: 30, loss: 0.000443245138740167
step: 40, loss: 0.0002249510434921831
step: 50, loss: 0.00046169955749064684
step: 60, loss: 0.0007141903624869883
step: 70, loss: 0.00020864907128270715
step: 80, loss: 0.0002377592900302261
step: 90, loss: 0.00042205804493278265
step: 100, loss: 0.0016757659614086151
step: 110, loss: 0.009700370021164417
step: 120, loss: 0.00023702230828348547
step: 130, loss: 0.0033940060529857874
step: 140, loss: 0.03564024716615677
step: 150, loss: 0.02369597554206848
step: 160, loss: 0.0024341177195310593
step: 170, loss: 0.06715569645166397
step: 180, loss: 0.007793682161718607
step: 190, loss: 0.017034141346812248
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.012957175262272358
step: 210, loss: 0.016357239335775375
step: 220, loss: 0.0020261872559785843
step: 230, loss: 0.00018657476175576448
/usr/local/lib/python3.7/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=0.977728285077951, f1=0.9764837625979844,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007505138055421412
step: 10, loss: 0.0003858654235955328
step: 20, loss: 0.05474752187728882
step: 30, loss: 0.00017649083747528493
step: 40, loss: 0.00024906941689550877
step: 50, loss: 0.0001924019306898117
step: 60, loss: 0.007871522568166256
step: 70, loss: 0.04101525992155075
step: 80, loss: 0.022572802379727364

```

```

step: 90, loss: 0.07587160170078278
step: 100, loss: 0.021740589290857315
step: 110, loss: 0.00036275279126130044
step: 120, loss: 0.0001301109150517732
step: 130, loss: 0.0010542593663558364
step: 140, loss: 0.0002748150727711618
step: 150, loss: 0.00026373000582680106
step: 160, loss: 0.0026004211977124214
step: 170, loss: 0.0004876042657997459
step: 180, loss: 0.021492116153240204
step: 190, loss: 0.04063574969768524
step: 200, loss: 0.0006855179672129452
step: 210, loss: 0.0024569567758589983
step: 220, loss: 0.00574579369276762
step: 230, loss: 0.009538508951663971
/usr/local/lib/python3.7/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=0.9777777777777777, f1=0.9732739420935412,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011323271319270134
step: 10, loss: 0.00023284034978132695
step: 20, loss: 0.0004553607141133398
step: 30, loss: 0.0005193978431634605
step: 40, loss: 0.00010867544915527105
step: 50, loss: 7.958558853715658e-05
step: 60, loss: 0.0027994865085929632
step: 70, loss: 0.00037448504008352757
step: 80, loss: 5.872127439943142e-05
step: 90, loss: 9.536505967844278e-05
step: 100, loss: 5.763748413301073e-05
step: 110, loss: 0.00033230966073460877
step: 120, loss: 6.16591569269076e-05
step: 130, loss: 0.00014604518946725875
step: 140, loss: 8.989129128167406e-05
step: 150, loss: 0.011093147099018097
step: 160, loss: 0.06474072486162186
step: 170, loss: 0.015190738253295422
step: 180, loss: 0.072641521692276
step: 190, loss: 0.0009698064532130957
step: 200, loss: 0.033833008259534836
step: 210, loss: 5.169623182155192e-05
step: 220, loss: 0.003871127963066101
step: 230, loss: 0.0006480291485786438
/usr/local/lib/python3.7/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=0.9764309764309763, f1=0.9786276715410572,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.9178339142818004e-05
step: 10, loss: 0.00023293413687497377
step: 20, loss: 0.0001673968945397064
step: 30, loss: 7.185059803305194e-05
step: 40, loss: 0.004231436178088188
step: 50, loss: 0.0006004883325658739
step: 60, loss: 0.003030693856999278
step: 70, loss: 0.0011070560431107879
step: 80, loss: 0.001008224906399846
step: 90, loss: 2.7793608751380816e-05
step: 100, loss: 0.0004266496398486197
step: 110, loss: 0.0007959711947478354
step: 120, loss: 0.0005954501684755087
step: 130, loss: 0.0008224673219956458
step: 140, loss: 0.0006018474814482033
step: 150, loss: 0.006748910993337631
step: 160, loss: 0.0001614090579096228
step: 170, loss: 0.0002175942063331604
step: 180, loss: 6.932447286089882e-05
step: 190, loss: 6.98227813700214e-05
step: 200, loss: 4.43750323029235e-05
step: 210, loss: 0.0037833370734006166
step: 220, loss: 0.01209774799644947
step: 230, loss: 5.99489503656514e-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=0.978675645342312, f1=0.9751131221719457,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.0141214412869886e-05
step: 10, loss: 0.0020203348249197006
step: 20, loss: 0.0003360837581567466
step: 30, loss: 5.033176421420649e-05
step: 40, loss: 0.007808235008269548
step: 50, loss: 0.00021543630282394588
step: 60, loss: 9.612626308808103e-05
step: 70, loss: 0.008340761996805668
step: 80, loss: 0.0002621329331304878

```

```

step: 90, loss: 0.0037684286944568157
step: 100, loss: 0.0019229378085583448
step: 110, loss: 0.0010363790206611156
step: 120, loss: 0.001116381143219769
step: 130, loss: 0.00019998452626168728
step: 140, loss: 0.0004088571004103869
step: 150, loss: 0.001927050412632525
step: 160, loss: 9.567171218805015e-05
step: 170, loss: 0.005590302869677544
step: 180, loss: 0.0005509807378984988
step: 190, loss: 0.00041721074376255274
step: 200, loss: 5.809206413687207e-05
step: 210, loss: 7.227940659504384e-05
step: 220, loss: 4.9317259254166856e-05
step: 230, loss: 0.0002313005388714373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9786276715410572, f1=0.9808773903262092,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.692706741276197e-05
step: 10, loss: 2.6925592464976944e-05
step: 20, loss: 8.21782523416914e-05
step: 30, loss: 4.6928151277825236e-05
step: 40, loss: 4.34523681178689e-05
step: 50, loss: 4.976428681402467e-05
step: 60, loss: 9.224157111020759e-05
step: 70, loss: 0.00025069701950997114
step: 80, loss: 4.011946657556109e-05
step: 90, loss: 8.578548295190558e-05
step: 100, loss: 3.737796578207053e-05
step: 110, loss: 0.0001412780547980219
step: 120, loss: 0.0015656538307666779
step: 130, loss: 4.92650069645606e-05
step: 140, loss: 0.024852976202964783
step: 150, loss: 0.08739501982927322
step: 160, loss: 5.6043369113467634e-05
step: 170, loss: 3.383718285476789e-05
step: 180, loss: 8.12620492069982e-05
step: 190, loss: 0.018120955675840378
step: 200, loss: 3.559690230758861e-05
step: 210, loss: 2.8816679332521744e-05
step: 220, loss: 2.9782644560327753e-05
step: 230, loss: 0.0021072220988571644
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9787709497206705,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.529914465616457e-05
step: 10, loss: 4.868388350587338e-05
step: 20, loss: 0.00018428699695505202
step: 30, loss: 0.001978818327188492
step: 40, loss: 0.00010132878378499299
step: 50, loss: 3.101183392573148e-05
step: 60, loss: 4.083069143234752e-05
step: 70, loss: 8.326130046043545e-05
step: 80, loss: 0.023242291063070297
step: 90, loss: 0.00035666528856381774
step: 100, loss: 3.0747487471671775e-05
step: 110, loss: 0.000199181042262353
step: 120, loss: 2.5882613044814207e-05
step: 130, loss: 3.991020639659837e-05
step: 140, loss: 0.0020656234119087458
step: 150, loss: 0.0372624509036541
step: 160, loss: 0.00012780290853697807
step: 170, loss: 0.01726827397942543
step: 180, loss: 4.298571730032563e-05
step: 190, loss: 2.5029446987900883e-05
step: 200, loss: 0.0008252472034655511
step: 210, loss: 2.10214639082551e-05
step: 220, loss: 0.0006213845917955041
step: 230, loss: 5.742997018387541e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9777777777777777, f1=0.9765886287625419,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014526712766382843
step: 10, loss: 0.0043494743295013905
step: 20, loss: 3.607795952120796e-05
step: 30, loss: 0.00014232499233912677
step: 40, loss: 0.00016306384350173175
step: 50, loss: 0.002519128378480673
step: 60, loss: 0.0007827311055734754
step: 70, loss: 0.0003259415098000318
step: 80, loss: 0.00011849199654534459

```

```

step: 90, loss: 5.0995193305425346e-05
step: 100, loss: 0.007795060984790325
step: 110, loss: 0.002199922688305378
step: 120, loss: 0.01666383258998394
step: 130, loss: 3.122459020232782e-05
step: 140, loss: 0.00014551296771969646
step: 150, loss: 0.00030263428925536573
step: 160, loss: 2.4266140826512128e-05
step: 170, loss: 5.861806494067423e-05
step: 180, loss: 0.0032070851884782314
step: 190, loss: 0.0003678657812997699
step: 200, loss: 0.0014230821980163455
step: 210, loss: 0.00021161034237593412
step: 220, loss: 0.00015664042439311743
step: 230, loss: 0.032081298530101776
/usr/local/lib/python3.7/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=0.9794988610478361, f1=0.9785310734463276,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014884456992149353
step: 10, loss: 0.0008423524559475482
step: 20, loss: 0.00031331469654105604
step: 30, loss: 0.00010718830162659287
step: 40, loss: 8.422827522736043e-05
step: 50, loss: 9.607362881070003e-05
step: 60, loss: 4.861156776314601e-05
step: 70, loss: 2.9536699003074318e-05
step: 80, loss: 3.639729038695805e-05
step: 90, loss: 0.00027899720589630306
step: 100, loss: 0.00021076883422210813
step: 110, loss: 0.04366595670580864
step: 120, loss: 0.03658502548933029
step: 130, loss: 0.0005641589523293078
step: 140, loss: 8.748171967454255e-05
step: 150, loss: 2.5148923668893985e-05
step: 160, loss: 0.00012186375533929095
step: 170, loss: 0.00033012841595336795
step: 180, loss: 0.011607036925852299
step: 190, loss: 3.5568911698646843e-05
step: 200, loss: 0.0001394519640598446
step: 210, loss: 0.00013396210852079093
step: 220, loss: 6.979445606702939e-05
step: 230, loss: 5.610302832792513e-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=0.9797752808988766, f1=0.9798206278026906,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3021877495921217e-05
step: 10, loss: 0.000571489566937089
step: 20, loss: 4.7818157327128574e-05
step: 30, loss: 2.9495991839212365e-05
step: 40, loss: 0.0002911607443820685
step: 50, loss: 0.00033745376276783645
step: 60, loss: 5.7101417041849345e-05
step: 70, loss: 0.00018171186093240976
step: 80, loss: 3.4572691220091656e-05
step: 90, loss: 2.4213881260948256e-05
step: 100, loss: 2.71714980044635e-05
step: 110, loss: 6.966407818254083e-05
step: 120, loss: 0.00012517468712758273
step: 130, loss: 5.769404015154578e-05
step: 140, loss: 9.3765978817828e-05
step: 150, loss: 0.00012859517300967127
step: 160, loss: 0.008047889918088913
step: 170, loss: 3.3670443372102454e-05
step: 180, loss: 6.520949682453647e-05
step: 190, loss: 2.8262831619940698e-05
step: 200, loss: 3.50938098563347e-05
step: 210, loss: 4.1609135223552585e-05
step: 220, loss: 3.5439239582046866e-05
step: 230, loss: 9.544281783746555e-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=0.9799107142857142, f1=0.9787709497206705,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.066966696176678e-05
step: 10, loss: 1.640969276195392e-05
step: 20, loss: 6.150231638457626e-05
step: 30, loss: 7.819347229087725e-05
step: 40, loss: 0.0002878784143831581
step: 50, loss: 0.0004801259783562273
step: 60, loss: 0.0001839497999753803
step: 70, loss: 0.0012119420571252704
step: 80, loss: 2.542409310990479e-05

```

```

step: 90, loss: 5.829031579196453e-05
step: 100, loss: 0.00013501591456588358
step: 110, loss: 0.0018160495674237609
step: 120, loss: 0.007287368644028902
step: 130, loss: 3.9845988794695586e-05
step: 140, loss: 3.173365257680416e-05
step: 150, loss: 9.672277519712225e-05
step: 160, loss: 0.0003824495943263173
step: 170, loss: 1.6767320630606264e-05
step: 180, loss: 0.00012545708159450442
step: 190, loss: 0.00014972765347920358
step: 200, loss: 0.000402468373067677
step: 210, loss: 2.5875326173263602e-05
step: 220, loss: 0.008084784261882305
step: 230, loss: 6.666833360213786e-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=0.98, f1=0.9788182831661093, best_f1=0.9819819819819819
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 [00:11, 206.12it/s]
load_f1 = 0.9831649831649831
real_f1 = 0.9832026875699889
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, 255.65it/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 O2: FP16 training with FP32 batchnorm and FP32

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

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

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

```
step: 0, loss: 0.6322342753410339
step: 10, loss: 0.5408704280853271
step: 20, loss: 0.523583710193634
step: 30, loss: 0.1490127146244049
step: 40, loss: 0.17368084192276
step: 50, loss: 0.18595919013023376
step: 60, loss: 0.06715504080057144
step: 70, loss: 0.1337631195783615
step: 80, loss: 0.04418506473302841
step: 90, loss: 0.2614085078239441
step: 100, loss: 0.03933422267436981
step: 110, loss: 0.03659733757376671
step: 120, loss: 0.12773193418979645
```

```

step: 130, loss: 0.06950389593839645
step: 140, loss: 0.21308594942092896
step: 150, loss: 0.04952492192387581
step: 160, loss: 0.10250748693943024
step: 170, loss: 0.19559890031814575
step: 180, loss: 0.10200154036283493
step: 190, loss: 0.03325225040316582
step: 200, loss: 0.19940772652626038
step: 210, loss: 0.09406846761703491
step: 220, loss: 0.2800471782684326
step: 230, loss: 0.17923583090305328
step: 240, loss: 0.10511860996484756
step: 250, loss: 0.024465041235089302
step: 260, loss: 0.13942429423332214
step: 270, loss: 0.027932170778512955
step: 280, loss: 0.08261004090309143
step: 290, loss: 0.08205541968345642
step: 300, loss: 0.08670815080404282
step: 310, loss: 0.24042752385139465
step: 320, loss: 0.14043736457824707
step: 330, loss: 0.05761094391345978
step: 340, loss: 0.07244287431240082
step: 350, loss: 0.22161749005317688
step: 360, loss: 0.13819853961467743
step: 370, loss: 0.12629184126853943
step: 380, loss: 0.019149092957377434
step: 390, loss: 0.05877387151122093
step: 400, loss: 0.2394690364599228
step: 410, loss: 0.04696466773748398
step: 420, loss: 0.058855608105659485
step: 430, loss: 0.17166511714458466
step: 440, loss: 0.014797236770391464
step: 450, loss: 0.008870395831763744
step: 460, loss: 0.01909269578754902
step: 470, loss: 0.14125855267047882
step: 480, loss: 0.09046563506126404
step: 490, loss: 0.21604043245315552
step: 500, loss: 0.11134421080350876
step: 510, loss: 0.07705353200435638
step: 520, loss: 0.0964888334274292
step: 530, loss: 0.0062089283019304276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.92772186642269, f1=0.923992673992674,  
best_f1=0.923992673992674  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.18741579353809357  
step: 10, loss: 0.03193799778819084  
step: 20, loss: 0.030968135222792625  
step: 30, loss: 0.14152728021144867  
step: 40, loss: 0.08025015145540237  
step: 50, loss: 0.19317074120044708  
step: 60, loss: 0.03094249963760376  
step: 70, loss: 0.035545628517866135  
step: 80, loss: 0.10316495597362518  
step: 90, loss: 0.07577692717313766  
step: 100, loss: 0.10903696715831757  
step: 110, loss: 0.019358310848474503  
step: 120, loss: 0.08123113214969635  
step: 130, loss: 0.15020935237407684  
step: 140, loss: 0.10068026930093765  
step: 150, loss: 0.04739644005894661  
step: 160, loss: 0.05438083037734032  
step: 170, loss: 0.16687148809432983  
step: 180, loss: 0.02023128978908062  
step: 190, loss: 0.044584907591342926  
step: 200, loss: 0.09051588922739029  
step: 210, loss: 0.15247783064842224  
step: 220, loss: 0.04429084435105324  
step: 230, loss: 0.012627234682440758  
step: 240, loss: 0.10429372638463974  
step: 250, loss: 0.011985398828983307  
step: 260, loss: 0.004522934556007385  
step: 270, loss: 0.2605387270450592  
step: 280, loss: 0.035486992448568344  
step: 290, loss: 0.03865816071629524  
step: 300, loss: 0.11652872711420059  
step: 310, loss: 0.07333706319332123  
step: 320, loss: 0.08177251368761063  
step: 330, loss: 0.06687607616186142  
step: 340, loss: 0.013469809666275978  
step: 350, loss: 0.08003512024879456  
step: 360, loss: 0.1662840098142624  
step: 370, loss: 0.15200942754745483  
step: 380, loss: 0.05153115838766098  
step: 390, loss: 0.11166876554489136  
step: 400, loss: 0.08494836837053299  
step: 410, loss: 0.06295786798000336  
step: 420, loss: 0.1950501948595047
```

```
step: 430, loss: 0.01576896756887436
step: 440, loss: 0.08162584900856018
step: 450, loss: 0.05522236227989197
step: 460, loss: 0.02621881105005741
step: 470, loss: 0.0931309387087822
step: 480, loss: 0.24021796882152557
step: 490, loss: 0.024218114092946053
step: 500, loss: 0.32096004486083984
step: 510, loss: 0.015589328482747078
step: 520, loss: 0.03632287308573723
step: 530, loss: 0.15255621075630188
/usr/local/lib/python3.7/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=0.9268738574040218, f1=0.9251824817518249,
best_f1=0.923992673992674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0441826730966568
step: 10, loss: 0.03131771832704544
step: 20, loss: 0.13764990866184235
step: 30, loss: 0.11220410466194153
step: 40, loss: 0.050907716155052185
step: 50, loss: 0.15568652749061584
step: 60, loss: 0.005260345060378313
step: 70, loss: 0.0094315679743886
step: 80, loss: 0.004941311199218035
step: 90, loss: 0.007868305779993534
step: 100, loss: 0.05615580081939697
step: 110, loss: 0.01150115579366684
step: 120, loss: 0.18542222678661346
step: 130, loss: 0.04202726110816002
step: 140, loss: 0.06769528239965439
step: 150, loss: 0.045699749141931534
step: 160, loss: 0.049739938229322433
step: 170, loss: 0.158523827791214
step: 180, loss: 0.061872292309999466
step: 190, loss: 0.008173956535756588
step: 200, loss: 0.03608530014753342
step: 210, loss: 0.10054024308919907
step: 220, loss: 0.050652723759412766
step: 230, loss: 0.11315387487411499
step: 240, loss: 0.010724533349275589
step: 250, loss: 0.04184827581048012
step: 260, loss: 0.11872739344835281
step: 270, loss: 0.006083154585212469
step: 280, loss: 0.14228668808937073
```

```

step: 290, loss: 0.006174887530505657
step: 300, loss: 0.09836077690124512
step: 310, loss: 0.08391545712947845
step: 320, loss: 0.019481206312775612
step: 330, loss: 0.008276775479316711
step: 340, loss: 0.009953660890460014
step: 350, loss: 0.009386810474097729
step: 360, loss: 0.026169266551733017
step: 370, loss: 0.005686056800186634
step: 380, loss: 0.018060801550745964
step: 390, loss: 0.008293497376143932
step: 400, loss: 0.0381668396294117
step: 410, loss: 0.026957081630825996
step: 420, loss: 0.02597564458847046
step: 430, loss: 0.07767129689455032
step: 440, loss: 0.024063993245363235
step: 450, loss: 0.09402237832546234
step: 460, loss: 0.03184078261256218
step: 470, loss: 0.05514726787805557
step: 480, loss: 0.007322478573769331
step: 490, loss: 0.01343369111418724
step: 500, loss: 0.002800171496346593
step: 510, loss: 0.011028800159692764
step: 520, loss: 0.10073144733905792
step: 530, loss: 0.057171713560819626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9353187529083293, f1=0.9273066169617894,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006281989626586437
step: 10, loss: 0.0027225385420024395
step: 20, loss: 0.0006790294428355992
step: 30, loss: 0.018199441954493523
step: 40, loss: 0.08359578996896744
step: 50, loss: 0.09942484647035599
step: 60, loss: 0.05751834437251091
step: 70, loss: 0.005214907228946686
step: 80, loss: 0.01380927488207817
step: 90, loss: 0.029539695009589195
step: 100, loss: 0.0048690298572182655

```

```
step: 110, loss: 0.023868869990110397
step: 120, loss: 0.03301266208291054
step: 130, loss: 0.01895071007311344
step: 140, loss: 0.004104114603251219
step: 150, loss: 0.002487856661900878
step: 160, loss: 0.0119104515761137
step: 170, loss: 0.003353266976773739
step: 180, loss: 0.015540062449872494
step: 190, loss: 0.020617786794900894
step: 200, loss: 0.013670647516846657
step: 210, loss: 0.14083324372768402
step: 220, loss: 0.01945408061146736
step: 230, loss: 0.07956930994987488
step: 240, loss: 0.011040345765650272
step: 250, loss: 0.05671041086316109
step: 260, loss: 0.006262488197535276
step: 270, loss: 0.0077774724923074245
step: 280, loss: 0.029507657513022423
step: 290, loss: 0.015324395149946213
step: 300, loss: 0.00026352997520007193
step: 310, loss: 0.0068644373677670956
step: 320, loss: 0.0352204404771328
step: 330, loss: 0.08082101494073868
step: 340, loss: 0.0028543462976813316
step: 350, loss: 0.03760666400194168
step: 360, loss: 0.00512770377099514
step: 370, loss: 0.009954886510968208
step: 380, loss: 0.0291168000549078
step: 390, loss: 0.050036389380693436
step: 400, loss: 0.0024203634820878506
step: 410, loss: 0.0025794277898967266
step: 420, loss: 0.008877637796103954
step: 430, loss: 0.24034643173217773
step: 440, loss: 0.011493087746202946
step: 450, loss: 0.0033056207466870546
step: 460, loss: 0.04608505591750145
step: 470, loss: 0.007756899110972881
step: 480, loss: 0.21438738703727722
step: 490, loss: 0.005909201223403215
step: 500, loss: 0.01734284870326519
step: 510, loss: 0.08859635144472122
step: 520, loss: 0.05940551310777664
step: 530, loss: 0.029966311529278755
/usr/local/lib/python3.7/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=0.9231464737793852, f1=0.9246823956442831,
best_f1=0.9273066169617894
```

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

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

```
step: 0, loss: 0.06552721560001373
```

```
step: 10, loss: 0.08055546134710312
```

```
step: 20, loss: 0.002584031317383051
```

```
step: 30, loss: 0.008968268521130085
```

```
step: 40, loss: 0.0095588443800807
```

```
step: 50, loss: 0.00863940455019474
```

```
step: 60, loss: 0.17101788520812988
```

```
step: 70, loss: 0.08944030851125717
```

```
step: 80, loss: 0.0014904835261404514
```

```
step: 90, loss: 0.007398639805614948
```

```
step: 100, loss: 0.024298040196299553
```

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

```
step: 110, loss: 0.0015928345965221524
```

```
step: 120, loss: 0.013186979107558727
```

```
step: 130, loss: 0.005009468644857407
```

```
step: 140, loss: 0.001338920439593494
```

```
step: 150, loss: 0.0027808877639472485
```

```
step: 160, loss: 0.002340702572837472
```

```
step: 170, loss: 0.03746198117733002
```

```
step: 180, loss: 0.0036231600679457188
```

```
step: 190, loss: 0.008374166674911976
```

```
step: 200, loss: 0.00197369116358459
```

```
step: 210, loss: 0.08988475799560547
```

```
step: 220, loss: 0.002155038993805647
```

```
step: 230, loss: 0.004746838007122278
```

```
step: 240, loss: 0.029375722631812096
```

```
step: 250, loss: 0.011292141862213612
```

```
step: 260, loss: 0.009541928768157959
```

```
step: 270, loss: 0.0006901574670337141
```

```
step: 280, loss: 0.015505513176321983
```

```
step: 290, loss: 0.12062451243400574
```

```
step: 300, loss: 0.015097645111382008
```

```
step: 310, loss: 0.009880187921226025
```

```
step: 320, loss: 0.09474422037601471
```

```
step: 330, loss: 0.05617875978350639
```

```
step: 340, loss: 0.0025574357714504004
```

```
step: 350, loss: 0.07778268307447433
```

```
step: 360, loss: 0.002325251465663314
```

```
step: 370, loss: 0.007440947461873293
```

```
step: 380, loss: 0.002786786062642932
```

```
step: 390, loss: 0.001908269478008151
```

```
step: 400, loss: 0.06112837791442871
```

```
step: 410, loss: 0.004828553646802902
```

```
step: 420, loss: 0.01616903580725193
```

```
step: 430, loss: 0.0044789682142436504
```



```
step: 440, loss: 0.021240651607513428
step: 450, loss: 0.0025660006795078516
step: 460, loss: 0.00352600053884089
step: 470, loss: 0.004376404453068972
step: 480, loss: 0.015406597405672073
step: 490, loss: 0.0009327676962129772
step: 500, loss: 0.035809434950351715
step: 510, loss: 0.01761043816804886
step: 520, loss: 0.0020885004196316004
step: 530, loss: 0.014144103974103928
/usr/local/lib/python3.7/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=0.9263947491795592, f1=0.9242990654205607,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011396723799407482
step: 10, loss: 0.0003842939913738519
step: 20, loss: 0.24195586144924164
step: 30, loss: 0.002217572880908847
step: 40, loss: 0.008290602825582027
step: 50, loss: 0.007047407794743776
step: 60, loss: 0.004112366121262312
step: 70, loss: 0.001399695174768567
step: 80, loss: 0.0005744108348153532
step: 90, loss: 0.04332095757126808
step: 100, loss: 0.003463148605078459
step: 110, loss: 0.0023967719171196222
step: 120, loss: 0.0015275543555617332
step: 130, loss: 0.012880127876996994
step: 140, loss: 0.04716102033853531
step: 150, loss: 0.0010325326584279537
step: 160, loss: 0.0010095973266288638
step: 170, loss: 0.04044920578598976
step: 180, loss: 0.0006842154543846846
step: 190, loss: 0.013887298293411732
step: 200, loss: 0.009918991476297379
step: 210, loss: 0.016008468344807625
step: 220, loss: 0.0037406287156045437
step: 230, loss: 0.005408292170614004
step: 240, loss: 0.0648561343550682
step: 250, loss: 0.0031741023994982243
step: 260, loss: 0.0015899627469480038
step: 270, loss: 0.1436314582824707
step: 280, loss: 0.06589912623167038
step: 290, loss: 0.001131078926846385
```

```

step: 300, loss: 0.02530484087765217
step: 310, loss: 0.002000723499804735
step: 320, loss: 0.004164023790508509
step: 330, loss: 0.01271990966051817
step: 340, loss: 0.15714266896247864
step: 350, loss: 0.024008123204112053
step: 360, loss: 0.004037526436150074
step: 370, loss: 0.00864609144628048
step: 380, loss: 0.0011871438473463058
step: 390, loss: 0.007738953456282616
step: 400, loss: 0.00029123135027475655
step: 410, loss: 0.0030174695421010256
step: 420, loss: 0.0010335989063605666
step: 430, loss: 0.0834420695900917
step: 440, loss: 0.10911819338798523
step: 450, loss: 0.00043778802501037717
step: 460, loss: 0.000617483863607049
step: 470, loss: 0.016385184600949287
step: 480, loss: 0.0071140252985060215
step: 490, loss: 0.004140689503401518
step: 500, loss: 0.00022630399325862527
step: 510, loss: 0.015991751104593277
step: 520, loss: 0.002099620411172509
step: 530, loss: 0.08393146097660065
/usr/local/lib/python3.7/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=0.924618320610687, f1=0.9234404536862004,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031169914291240275
step: 10, loss: 0.010283399373292923
step: 20, loss: 0.001684085582382977
step: 30, loss: 0.0008984761661849916
step: 40, loss: 0.000498548208270222
step: 50, loss: 0.0064149415120482445
step: 60, loss: 0.0485406257212162
step: 70, loss: 0.03753560036420822
step: 80, loss: 0.009361266158521175
step: 90, loss: 0.0001643301802687347
step: 100, loss: 0.0010167600121349096
step: 110, loss: 0.0011050673201680183
step: 120, loss: 0.0002119772252626717
step: 130, loss: 0.003117576241493225
step: 140, loss: 0.00024349008162971586
step: 150, loss: 0.0174198467284441

```

```

step: 160, loss: 0.0007258100667968392
step: 170, loss: 0.007008664309978485
step: 180, loss: 0.02053200453519821
step: 190, loss: 0.009550072252750397
step: 200, loss: 0.0008869237499311566
step: 210, loss: 0.016557075083255768
step: 220, loss: 0.0005868254229426384
step: 230, loss: 0.005909034516662359
step: 240, loss: 0.01658540405333042
step: 250, loss: 0.01894763670861721
step: 260, loss: 0.0009835308883339167
step: 270, loss: 0.1291879266500473
step: 280, loss: 0.0007495644385926425
step: 290, loss: 0.0010711801005527377
step: 300, loss: 0.002543291077017784
step: 310, loss: 0.001371196936815977
step: 320, loss: 0.00011956487287534401
step: 330, loss: 0.0007326148333959281
step: 340, loss: 0.015881575644016266
step: 350, loss: 0.014363614842295647
step: 360, loss: 0.0005018475349061191
step: 370, loss: 0.00035022158408537507
step: 380, loss: 0.03287065401673317
step: 390, loss: 0.0036975329276174307
step: 400, loss: 0.008704678155481815
step: 410, loss: 0.0063106948509812355
step: 420, loss: 0.028039639815688133
step: 430, loss: 0.00027445924933999777
step: 440, loss: 0.00043252581963315606
step: 450, loss: 0.0017018248327076435
step: 460, loss: 0.00016534203314222395
step: 470, loss: 0.014156284742057323
step: 480, loss: 0.10643943399190903
step: 490, loss: 0.004011459648609161
step: 500, loss: 0.0007829935639165342
step: 510, loss: 0.0009510979289188981
step: 520, loss: 0.11424171924591064
step: 530, loss: 0.18629665672779083
/usr/local/lib/python3.7/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=0.9178470254957507, f1=0.9111933395004626,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006069974973797798
step: 10, loss: 0.021028082817792892

```

step: 20, loss: 0.0015337337972596288
step: 30, loss: 0.002116419840604067
step: 40, loss: 0.00021677042241208255
step: 50, loss: 0.0053200083784759045
step: 60, loss: 0.0012443923624232411
step: 70, loss: 0.0001550250017317012
step: 80, loss: 0.018107039853930473
step: 90, loss: 0.0002697952149901539
step: 100, loss: 0.0007751962402835488
step: 110, loss: 0.024284930899739265
step: 120, loss: 0.0012684524990618229
step: 130, loss: 0.00023085673456080258
step: 140, loss: 0.005839817691594362
step: 150, loss: 0.002996866125613451
step: 160, loss: 0.013097675517201424
step: 170, loss: 0.001117425737902522
step: 180, loss: 0.0025297889951616526
step: 190, loss: 0.014243213459849358
step: 200, loss: 0.0036141551099717617
step: 210, loss: 0.0007524169632233679
step: 220, loss: 0.0007020970224402845
step: 230, loss: 0.00163585739210248
step: 240, loss: 0.0011147791519761086
step: 250, loss: 0.00012311417958699167
step: 260, loss: 0.00033421628177165985
step: 270, loss: 0.007759368512779474
step: 280, loss: 0.00587490014731884
step: 290, loss: 0.00011610790534177795
step: 300, loss: 0.00013435517030302435
step: 310, loss: 0.005830511916428804
step: 320, loss: 0.014820123091340065
step: 330, loss: 0.0033891331404447556
step: 340, loss: 0.0005955698434263468
step: 350, loss: 4.954752148478292e-05
step: 360, loss: 0.0008657635189592838
step: 370, loss: 0.00021888338960707188
step: 380, loss: 0.0007376870489679277
step: 390, loss: 0.11908270418643951
step: 400, loss: 0.00012988634989596903
step: 410, loss: 0.00021460592688526958
step: 420, loss: 0.0018212919821962714
step: 430, loss: 0.01860579289495945
step: 440, loss: 0.002963392762467265
step: 450, loss: 0.0007605324499309063
step: 460, loss: 0.0012300023809075356
step: 470, loss: 0.0005581468576565385
step: 480, loss: 0.0025270995683968067
step: 490, loss: 0.0012828679755330086

```

step: 500, loss: 0.002538844244554639
step: 510, loss: 0.0005600863951258361
step: 520, loss: 0.0011736652813851833
step: 530, loss: 0.002243548398837447
/usr/local/lib/python3.7/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=0.9247211895910781, f1=0.9270544783010157,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007611307664774358
step: 10, loss: 0.0018057284178212285
step: 20, loss: 0.002756665460765362
step: 30, loss: 0.000695208553224802
step: 40, loss: 0.0002595694677438587
step: 50, loss: 0.0009759735548868775
step: 60, loss: 4.942653686157428e-05
step: 70, loss: 0.0004642940766643733
step: 80, loss: 0.0017070862231776118
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 90, loss: 0.0001529146684333682
step: 100, loss: 0.0017023079562932253
step: 110, loss: 0.0001562800316605717
step: 120, loss: 0.001937009976245463
step: 130, loss: 0.02670702524483204
step: 140, loss: 0.05635204538702965
step: 150, loss: 7.427267701132223e-05
step: 160, loss: 0.0037294584326446056
step: 170, loss: 0.013267877511680126
step: 180, loss: 0.00042188374209217727
step: 190, loss: 0.000311357929604128
step: 200, loss: 0.0003205654793418944
step: 210, loss: 0.00019440613687038422
step: 220, loss: 0.007061345037072897
step: 230, loss: 0.00015062035527080297
step: 240, loss: 0.1413906365633011
step: 250, loss: 0.0016124008689075708
step: 260, loss: 0.026845822110772133
step: 270, loss: 3.8081696402514353e-05
step: 280, loss: 0.014518075622618198
step: 290, loss: 0.048134904354810715
step: 300, loss: 0.00033126151538453996
step: 310, loss: 0.007221641950309277
step: 320, loss: 0.001232979353517294
step: 330, loss: 0.024264028295874596
step: 340, loss: 0.001116272178478539

```

```

step: 350, loss: 0.001178797334432602
step: 360, loss: 0.00013610887981485575
step: 370, loss: 0.0041011618450284
step: 380, loss: 0.001449335366487503
step: 390, loss: 0.0004500871291384101
step: 400, loss: 0.007555874530225992
step: 410, loss: 0.0016465706285089254
step: 420, loss: 0.00013570925511885434
step: 430, loss: 0.0002992751251440495
step: 440, loss: 0.0017538603860884905
step: 450, loss: 0.00043697957880795
step: 460, loss: 0.00044999580131843686
step: 470, loss: 0.0001290318905375898
step: 480, loss: 0.0008040433749556541
step: 490, loss: 0.004260864574462175
step: 500, loss: 0.0009129149257205427
step: 510, loss: 0.0017464299453422427
step: 520, loss: 0.0003665339609142393
step: 530, loss: 0.0002987234038300812
/usr/local/lib/python3.7/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=0.9233627496516488, f1=0.923992673992674,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002370211877860129
step: 10, loss: 0.00015046342741698027
step: 20, loss: 0.0001416628365404904
step: 30, loss: 4.8325484385713935e-05
step: 40, loss: 0.006672816816717386
step: 50, loss: 0.0004019698826596141
step: 60, loss: 9.159769979305565e-05
step: 70, loss: 0.0005965133314020932
step: 80, loss: 0.0011847903952002525
step: 90, loss: 0.00201031775213778
step: 100, loss: 0.0001068591300281696
step: 110, loss: 0.00015900017751846462
step: 120, loss: 9.210191637976095e-05
step: 130, loss: 0.00010701100836740807
step: 140, loss: 0.004450415261089802
step: 150, loss: 0.003194632474333048
step: 160, loss: 0.0028831639792770147
step: 170, loss: 0.0002711694687604904
step: 180, loss: 0.0020988646429032087
step: 190, loss: 0.0004730326181743294
step: 200, loss: 0.0018034002278000116

```

```

step: 210, loss: 0.0029832343570888042
step: 220, loss: 0.001314705703407526
step: 230, loss: 0.003892711130902171
step: 240, loss: 0.0004599791718646884
step: 250, loss: 0.062134455889463425
step: 260, loss: 0.00092978187603876
step: 270, loss: 0.00012802882702089846
step: 280, loss: 0.001449195435270667
step: 290, loss: 0.0006390062626451254
step: 300, loss: 8.036869985517114e-05
step: 310, loss: 0.0002004581328947097
step: 320, loss: 0.004559352993965149
step: 330, loss: 0.004852449521422386
step: 340, loss: 0.0003251806483604014
step: 350, loss: 0.022897906601428986
step: 360, loss: 0.014567120932042599
step: 370, loss: 0.0030197049491107464
step: 380, loss: 0.0015855649253353477
step: 390, loss: 0.00034738730755634606
step: 400, loss: 0.0024525783956050873
step: 410, loss: 0.004069699440151453
step: 420, loss: 0.00026663931203074753
step: 430, loss: 0.00011809538409579545
step: 440, loss: 0.007883870042860508
step: 450, loss: 8.055928628891706e-05
step: 460, loss: 0.0001110490266000852
step: 470, loss: 0.0001510552247054875
step: 480, loss: 2.3342281565419398e-05
step: 490, loss: 0.00026279629673808813
step: 500, loss: 0.003530255751684308
step: 510, loss: 0.0001862456410890445
step: 520, loss: 0.00018313944747205824
step: 530, loss: 0.0010759696597233415
/usr/local/lib/python3.7/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=0.9251059821008009, f1=0.925891181988743,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004691213835030794
step: 10, loss: 0.010594181716442108
step: 20, loss: 0.00043823086889460683
step: 30, loss: 0.0006772693595848978
step: 40, loss: 0.05836055055260658
step: 50, loss: 0.0015192842110991478
step: 60, loss: 0.02075786516070366

```

step: 70, loss: 0.0002280693151988089
step: 80, loss: 0.003908604383468628
step: 90, loss: 5.3101361118024215e-05
step: 100, loss: 0.0018792742630466819
step: 110, loss: 0.0007329937652684748
step: 120, loss: 0.00046343295252881944
step: 130, loss: 0.005455214995890856
step: 140, loss: 0.004025232512503862
step: 150, loss: 7.137224019970745e-05
step: 160, loss: 0.00016604815027676523
step: 170, loss: 8.742364298086613e-05
step: 180, loss: 0.02160881645977497
step: 190, loss: 0.00191867642570287
step: 200, loss: 0.003695595543831587
step: 210, loss: 0.0007218162063509226
step: 220, loss: 0.001233998453244567
step: 230, loss: 0.0002964069426525384
step: 240, loss: 0.0009064035839401186
step: 250, loss: 6.347743328660727e-05
step: 260, loss: 0.00010244065924780443
step: 270, loss: 0.0026955304201692343
step: 280, loss: 0.00020181421132292598
step: 290, loss: 0.003526706248521805
step: 300, loss: 0.004871679469943047
step: 310, loss: 0.011182836256921291
step: 320, loss: 8.768749830778688e-05
step: 330, loss: 0.0004986825515516102
step: 340, loss: 0.16435886919498444
step: 350, loss: 0.0005604855250567198
step: 360, loss: 0.002063812455162406
step: 370, loss: 0.0022219240199774504
step: 380, loss: 0.0001482973893871531
step: 390, loss: 0.00020173768280074
step: 400, loss: 0.0011890018358826637
step: 410, loss: 8.836227789288387e-05
step: 420, loss: 0.0001268740015802905
step: 430, loss: 3.462496169959195e-05
step: 440, loss: 0.001476971199735999
step: 450, loss: 0.00011536334204720333
step: 460, loss: 0.001494944328442216
step: 470, loss: 0.00022059044567868114
step: 480, loss: 0.0015208958648145199
step: 490, loss: 0.0010819482849910855
step: 500, loss: 0.003576661692932248
step: 510, loss: 0.00036646644002757967
step: 520, loss: 1.7646427295403555e-05
step: 530, loss: 8.498256647726521e-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=0.9279999999999999, f1=0.9296037296037295,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.2978270357707515e-05
step: 10, loss: 4.0420040022581816e-05
step: 20, loss: 0.00010819955787155777
step: 30, loss: 0.0007759088766761124
step: 40, loss: 0.0001839964825194329
step: 50, loss: 0.04978008568286896
step: 60, loss: 0.0014046597061678767
step: 70, loss: 0.001565490267239511
step: 80, loss: 0.00023644852626603097
step: 90, loss: 0.00012131755647715181
step: 100, loss: 0.00017519001266919076
step: 110, loss: 0.0002973974624183029
step: 120, loss: 0.0004714673850685358
step: 130, loss: 0.0012158085592091084
step: 140, loss: 1.880119634734001e-05
step: 150, loss: 4.7683723096270114e-05
step: 160, loss: 2.5755769456736743e-05
step: 170, loss: 0.00017819782078731805
step: 180, loss: 6.908526120241731e-05
step: 190, loss: 6.609653064515442e-05
step: 200, loss: 4.973252725903876e-05
step: 210, loss: 6.108388333814219e-05
step: 220, loss: 0.008433660492300987
step: 230, loss: 1.7854783436632715e-05
step: 240, loss: 0.008126065135002136
step: 250, loss: 2.3058302758727223e-05
step: 260, loss: 5.1883998821722344e-05
step: 270, loss: 0.00550859747454524
step: 280, loss: 0.0005938881658948958
step: 290, loss: 0.005603653844445944
step: 300, loss: 0.09377526491880417
step: 310, loss: 0.0007598432712256908
step: 320, loss: 3.024237412319053e-05
step: 330, loss: 7.136902422644198e-05
step: 340, loss: 0.00011924338468816131
step: 350, loss: 0.00018772676412481815
step: 360, loss: 0.004514878615736961
step: 370, loss: 0.3295959532260895
step: 380, loss: 0.00030403019627556205
step: 390, loss: 0.01745247095823288
step: 400, loss: 5.3141651733312756e-05

```

```

step: 410, loss: 0.00027573155239224434
step: 420, loss: 0.024752870202064514
step: 430, loss: 0.004485808778554201
step: 440, loss: 9.63168204179965e-05
step: 450, loss: 0.00015199987683445215
step: 460, loss: 6.193441367940977e-05
step: 470, loss: 0.0806107223033905
step: 480, loss: 0.00018795568030327559
step: 490, loss: 0.0007843024795874953
step: 500, loss: 8.345155220013112e-05
step: 510, loss: 0.0028137427289038897
step: 520, loss: 0.000301576656056568
step: 530, loss: 2.9003864256083034e-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=0.924020764511562, f1=0.9256661991584852,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000470706116175279
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.0017225549090653658
step: 20, loss: 0.00015944623737595975
step: 30, loss: 0.0003991667181253433
step: 40, loss: 4.0115231968229637e-05
step: 50, loss: 0.005064442288130522
step: 60, loss: 0.0007603391422890127
step: 70, loss: 0.00029997562523931265
step: 80, loss: 0.0001318665308644995
step: 90, loss: 0.037893764674663544
step: 100, loss: 0.00027594377752393484
step: 110, loss: 5.385743861552328e-05
step: 120, loss: 0.0002875733480323106
step: 130, loss: 0.00028979225317016244
step: 140, loss: 0.00020767253590747714
step: 150, loss: 0.0003457533603068441
step: 160, loss: 0.07452693581581116
step: 170, loss: 0.0002612207317724824
step: 180, loss: 2.906763256760314e-05
step: 190, loss: 0.0018670862773433328
step: 200, loss: 8.14655504655093e-05
step: 210, loss: 0.00017356194439344108
step: 220, loss: 0.00030821646214462817
step: 230, loss: 0.0003135166480205953
step: 240, loss: 0.0003976882726419717
step: 250, loss: 0.00015571160474792123

```

```

step: 260, loss: 0.0008520273258909583
step: 270, loss: 3.704003393067978e-05
step: 280, loss: 6.37486664345488e-05
step: 290, loss: 8.542381692677736e-05
step: 300, loss: 0.0001514780306024477
step: 310, loss: 0.004474738147109747
step: 320, loss: 0.01249807421118021
step: 330, loss: 0.0033873948268592358
step: 340, loss: 0.00019889218674506992
step: 350, loss: 0.0003093962150160223
step: 360, loss: 6.572713755303994e-05
step: 370, loss: 0.00012080020678695291
step: 380, loss: 0.00045204575872048736
step: 390, loss: 0.001572229783050716
step: 400, loss: 0.0001613130298210308
step: 410, loss: 0.001028162776492536
step: 420, loss: 7.079893111949787e-05
step: 430, loss: 0.0013663740828633308
step: 440, loss: 0.001630643499083817
step: 450, loss: 0.006842336617410183
step: 460, loss: 0.007156752049922943
step: 470, loss: 0.00027096335543319583
step: 480, loss: 6.19588972767815e-05
step: 490, loss: 6.666514673270285e-05
step: 500, loss: 4.8294063162757084e-05
step: 510, loss: 0.00029093417106196284
step: 520, loss: 0.003515374381095171
step: 530, loss: 4.514653846854344e-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=0.9235880398671096, f1=0.9262564584311883,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000506674696225673
step: 10, loss: 0.00020781852072104812
step: 20, loss: 0.0004485551326069981
step: 30, loss: 3.95294810004998e-05
step: 40, loss: 0.00184946833178401
step: 50, loss: 0.0009200516506098211
step: 60, loss: 0.0011141477152705193
step: 70, loss: 3.8756617868784815e-05
step: 80, loss: 0.00022810566588304937
step: 90, loss: 0.0002603710163384676
step: 100, loss: 0.002347462810575962
step: 110, loss: 0.00015681145305279642

```

```
step: 120, loss: 0.0001551572058815509
step: 130, loss: 0.016474341973662376
step: 140, loss: 0.008556803688406944
step: 150, loss: 6.222486990736797e-05
step: 160, loss: 5.077393871033564e-05
step: 170, loss: 0.001421546097844839
step: 180, loss: 2.6619989512255415e-05
step: 190, loss: 0.00010893047146964818
step: 200, loss: 0.002607255708426237
step: 210, loss: 0.0001302546152146533
step: 220, loss: 4.865380105911754e-05
step: 230, loss: 0.018987605348229408
step: 240, loss: 0.00016796017007436603
step: 250, loss: 0.00010934309830190614
step: 260, loss: 0.0023119794204831123
step: 270, loss: 6.551338447025046e-05
step: 280, loss: 6.108253728598356e-05
step: 290, loss: 0.0008294066064991057
step: 300, loss: 0.0001624959404580295
step: 310, loss: 0.003957983106374741
step: 320, loss: 0.0003006841870956123
step: 330, loss: 0.00021098021534271538
step: 340, loss: 0.05189394950866699
step: 350, loss: 0.001737135462462902
step: 360, loss: 0.00045470992336049676
step: 370, loss: 7.87892349762842e-05
step: 380, loss: 9.796453377930447e-05
step: 390, loss: 0.015381338074803352
step: 400, loss: 0.00012614102161023766
step: 410, loss: 7.005969382589683e-05
step: 420, loss: 3.9169026422314346e-05
step: 430, loss: 0.002239689463749528
step: 440, loss: 0.00014116572856437415
step: 450, loss: 0.00015059176075737923
step: 460, loss: 7.119214569684118e-05
step: 470, loss: 0.0011633670656010509
step: 480, loss: 0.0001111097153625451
step: 490, loss: 3.5024306271225214e-05
step: 500, loss: 0.00013855966972187161
step: 510, loss: 0.0006861527217552066
step: 520, loss: 5.3344785555964336e-05
step: 530, loss: 5.25333707628306e-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=0.9202226345083488, f1=0.9247706422018349,
best_f1=0.9273066169617894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 1.6852984117576852e-05
step: 10, loss: 1.429372150596464e-05
step: 20, loss: 0.00014547089813277125
step: 30, loss: 7.193500641733408e-05
step: 40, loss: 0.004681602586060762
step: 50, loss: 0.0037333837244659662
step: 60, loss: 2.095800118695479e-05
step: 70, loss: 0.0007054278394207358
step: 80, loss: 4.034408266306855e-05
step: 90, loss: 3.004320024047047e-05
step: 100, loss: 0.0001027195030474104
step: 110, loss: 0.0032089308369904757
step: 120, loss: 0.031122397631406784
step: 130, loss: 0.18771006166934967
step: 140, loss: 0.00016600980598013848
step: 150, loss: 5.674261410604231e-05
step: 160, loss: 0.00016258865070994943
step: 170, loss: 5.517115641850978e-05
step: 180, loss: 0.00036490033380687237
step: 190, loss: 0.0018414832884445786
step: 200, loss: 0.04084758833050728
step: 210, loss: 0.0006854118546471
step: 220, loss: 5.6837176089175045e-05
step: 230, loss: 0.00011865880514960736
step: 240, loss: 5.450102617032826e-05
step: 250, loss: 3.6990008084103465e-05
step: 260, loss: 4.274760794942267e-05
step: 270, loss: 2.7918886189581826e-05
step: 280, loss: 3.55015326931607e-05
step: 290, loss: 1.754952063492965e-05
step: 300, loss: 1.9583469111239538e-05
step: 310, loss: 6.470639345934615e-05
step: 320, loss: 0.0002427146100671962
step: 330, loss: 2.0335814042482525e-05
step: 340, loss: 3.129323886241764e-05
step: 350, loss: 2.988901724165771e-05
step: 360, loss: 0.0002187026693718508
step: 370, loss: 3.4548436815384775e-05
step: 380, loss: 0.0007116995984688401
step: 390, loss: 0.002041775966063142
step: 400, loss: 0.00010049864067696035
step: 410, loss: 5.5362965213134885e-05
step: 420, loss: 0.00010748136264737695
step: 430, loss: 0.0008655879646539688
step: 440, loss: 6.407079490600154e-05
step: 450, loss: 1.7683545593172312e-05
```

```

step: 460, loss: 5.3089002904016525e-05
step: 470, loss: 0.0004505760734900832
step: 480, loss: 5.596923074335791e-05
step: 490, loss: 1.57465910888277e-05
step: 500, loss: 0.0427996963262558
step: 510, loss: 0.0006114896386861801
step: 520, loss: 0.00010259923874400556
step: 530, loss: 0.0003505299100652337
/usr/local/lib/python3.7/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=0.9251893939393939, f1=0.9240150093808631,
best_f1=0.9273066169617894
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 [00:20, 275.07it/s]
load_f1 = 0.9395348837209302
real_f1 = 0.9406307977736549
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
4398it [00:16, 267.29it/s]
```

4.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5627313256263733
step: 10, loss: 0.3636850118637085
step: 20, loss: 0.36867907643318176
step: 30, loss: 0.3171383738517761
step: 40, loss: 0.17773061990737915
step: 50, loss: 0.42412233352661133
step: 60, loss: 0.2381936013698578
step: 70, loss: 0.19261719286441803
step: 80, loss: 0.21679040789604187
step: 90, loss: 0.39025482535362244
step: 100, loss: 0.35806670784950256
step: 110, loss: 0.2342950850725174
step: 120, loss: 0.2790891230106354
step: 130, loss: 0.2325652837753296
step: 140, loss: 0.16302460432052612
step: 150, loss: 0.2714042365550995
step: 160, loss: 0.34730198979377747
step: 170, loss: 0.2665281295776367
step: 180, loss: 0.17581890523433685

```



```

step: 190, loss: 0.24125449359416962
step: 200, loss: 0.33124417066574097
step: 210, loss: 0.1888381987810135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5719298245614035, f1=0.6082474226804123,
best_f1=0.6082474226804123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11632522195577621
step: 10, loss: 0.19165131449699402
step: 20, loss: 0.17832064628601074
step: 30, loss: 0.19204361736774445
step: 40, loss: 0.2147732675075531
step: 50, loss: 0.1836603730916977
step: 60, loss: 0.37626558542251587
step: 70, loss: 0.11170294135808945
step: 80, loss: 0.20137391984462738
step: 90, loss: 0.09013444930315018
step: 100, loss: 0.02407083846628666
step: 110, loss: 0.12079711258411407
step: 120, loss: 0.18295852839946747
step: 130, loss: 0.03479720652103424
step: 140, loss: 0.14249208569526672
step: 150, loss: 0.24641817808151245
step: 160, loss: 0.19672951102256775
step: 170, loss: 0.20949013531208038
step: 180, loss: 0.1504782736301422
step: 190, loss: 0.21385183930397034
step: 200, loss: 0.05945597216486931
step: 210, loss: 0.1727837175130844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5955056179775281, f1=0.598848368522073,
best_f1=0.598848368522073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07283313572406769
step: 10, loss: 0.15944956243038177
step: 20, loss: 0.13192914426326752
step: 30, loss: 0.18072094023227692
step: 40, loss: 0.09981321543455124
step: 50, loss: 0.044434867799282074
step: 60, loss: 0.14078868925571442
step: 70, loss: 0.18338023126125336
step: 80, loss: 0.1837131828069687
step: 90, loss: 0.053550902754068375
step: 100, loss: 0.2548820376396179
step: 110, loss: 0.15040093660354614
step: 120, loss: 0.11738014221191406
step: 130, loss: 0.10153563320636749
step: 140, loss: 0.13115614652633667
step: 150, loss: 0.19873987138271332
step: 160, loss: 0.09167853742837906
step: 170, loss: 0.18252359330654144
step: 180, loss: 0.11461610347032547
step: 190, loss: 0.14666791260242462
step: 200, loss: 0.049963776022195816
step: 210, loss: 0.10398021340370178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6459627329192545, f1=0.6419753086419754,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13195273280143738
step: 10, loss: 0.04108063876628876
step: 20, loss: 0.20995505154132843
step: 30, loss: 0.12753646075725555
step: 40, loss: 0.02228063903748989
step: 50, loss: 0.2024925947189331
step: 60, loss: 0.18741635978221893
step: 70, loss: 0.2945934236049652
step: 80, loss: 0.12492763996124268
step: 90, loss: 0.028382305055856705
step: 100, loss: 0.2561679482460022
step: 110, loss: 0.32127100229263306
step: 120, loss: 0.23987482488155365

```

```

step: 130, loss: 0.1893261969089508
step: 140, loss: 0.1263992339372635
step: 150, loss: 0.0391581691801548
step: 160, loss: 0.1050899475812912
step: 170, loss: 0.146198570728302
step: 180, loss: 0.34382691979408264
step: 190, loss: 0.10990410298109055
step: 200, loss: 0.058949414640665054
step: 210, loss: 0.11207396537065506
/usr/local/lib/python3.7/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=0.5999999999999999, f1=0.6259842519685039,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09496463090181351
step: 10, loss: 0.08974441885948181
step: 20, loss: 0.21965420246124268
step: 30, loss: 0.10705877840518951
step: 40, loss: 0.022767236456274986
step: 50, loss: 0.19138777256011963
step: 60, loss: 0.16287483274936676
step: 70, loss: 0.1071845144033432
step: 80, loss: 0.06869249790906906
step: 90, loss: 0.16836583614349365
step: 100, loss: 0.01480349525809288
step: 110, loss: 0.10538613051176071
step: 120, loss: 0.180852010846138
step: 130, loss: 0.11777577549219131
step: 140, loss: 0.13320188224315643
step: 150, loss: 0.0382142998278141
step: 160, loss: 0.19265957176685333
step: 170, loss: 0.17808300256729126
step: 180, loss: 0.06284131854772568
step: 190, loss: 0.020699240267276764
step: 200, loss: 0.06842151284217834
step: 210, loss: 0.017668914049863815
/usr/local/lib/python3.7/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=0.5969498910675382, f1=0.611353711790393,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11723175644874573

```

```

step: 10, loss: 0.060088541358709335
step: 20, loss: 0.020140578970313072
step: 30, loss: 0.001891078194603324
step: 40, loss: 0.06721244752407074
step: 50, loss: 0.029240857809782028
step: 60, loss: 0.19988618791103363
step: 70, loss: 0.03422854095697403
step: 80, loss: 0.14502036571502686
step: 90, loss: 0.14777928590774536
step: 100, loss: 0.029405958950519562
step: 110, loss: 0.07534154504537582
step: 120, loss: 0.0666266530752182
step: 130, loss: 0.08941447734832764
step: 140, loss: 0.11189819872379303
step: 150, loss: 0.11905359476804733
step: 160, loss: 0.006318338215351105
step: 170, loss: 0.03336084634065628
step: 180, loss: 0.011646774597465992
step: 190, loss: 0.2904628813266754
step: 200, loss: 0.002397349802777171
step: 210, loss: 0.04880882054567337
/usr/local/lib/python3.7/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=0.6138211382113822, f1=0.598360655737705,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031162869185209274
step: 10, loss: 0.03560047596693039
step: 20, loss: 0.05501590296626091
step: 30, loss: 0.03509599342942238
step: 40, loss: 0.04214245826005936
step: 50, loss: 0.11953133344650269
step: 60, loss: 0.09149319678544998
step: 70, loss: 0.020939458161592484
step: 80, loss: 0.09098031371831894
step: 90, loss: 0.11441361159086227
step: 100, loss: 0.01297060027718544
step: 110, loss: 0.09997212886810303
step: 120, loss: 0.11763186752796173
step: 130, loss: 0.045648243278265
step: 140, loss: 0.05159392207860947
step: 150, loss: 0.027293652296066284
step: 160, loss: 0.018603915348649025
step: 170, loss: 0.009838255122303963
step: 180, loss: 0.024453474208712578

```

```

step: 190, loss: 0.18431676924228668
step: 200, loss: 0.023902345448732376
step: 210, loss: 0.0508066788315773
/usr/local/lib/python3.7/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=0.5906313645621182, f1=0.5978494623655914,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013351015746593475
step: 10, loss: 0.10731825977563858
step: 20, loss: 0.0024179001338779926
step: 30, loss: 0.04445340484380722
step: 40, loss: 0.0030142266768962145
step: 50, loss: 0.010832877829670906
step: 60, loss: 0.06737146526575089
step: 70, loss: 0.10593615472316742
step: 80, loss: 0.026981748640537262
step: 90, loss: 0.012885301373898983
step: 100, loss: 0.0973348543047905
step: 110, loss: 0.08062287420034409
step: 120, loss: 0.04261203482747078
step: 130, loss: 0.0021453818771988153
step: 140, loss: 0.06673210859298706
step: 150, loss: 0.0198410265147686
step: 160, loss: 0.030985843390226364
step: 170, loss: 0.05850386619567871
step: 180, loss: 0.07343168556690216
step: 190, loss: 0.011654342524707317
step: 200, loss: 0.018703462556004524
step: 210, loss: 0.22479870915412903
/usr/local/lib/python3.7/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=0.6031042128603104, f1=0.6132723112128146,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008775776252150536
step: 10, loss: 0.05210297182202339
step: 20, loss: 0.029559364542365074
step: 30, loss: 0.0253071878105402
step: 40, loss: 0.199273943901062
step: 50, loss: 0.0484335720539093
step: 60, loss: 0.08868417143821716

```

```

step: 70, loss: 0.01088655274361372
step: 80, loss: 0.016815800219774246
step: 90, loss: 0.003656730055809021
step: 100, loss: 0.009903131984174252
step: 110, loss: 0.042206063866615295
step: 120, loss: 0.02244407869875431
step: 130, loss: 0.05750494822859764
step: 140, loss: 0.044609054923057556
step: 150, loss: 0.18058113753795624
step: 160, loss: 0.015905214473605156
step: 170, loss: 0.007599723991006613
step: 180, loss: 0.013260343112051487
step: 190, loss: 0.001918022520840168
step: 200, loss: 0.0058121709153056145
step: 210, loss: 0.008510168641805649
/usr/local/lib/python3.7/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=0.5891783567134268, f1=0.5849462365591398,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10783641040325165
step: 10, loss: 0.004258967470377684
step: 20, loss: 0.006707071326673031
step: 30, loss: 0.1612420380115509
step: 40, loss: 0.002313343109562993
step: 50, loss: 0.021190805360674858
step: 60, loss: 0.011554659344255924
step: 70, loss: 0.047579552978277206
step: 80, loss: 0.014581447467207909
step: 90, loss: 0.04116307944059372
step: 100, loss: 0.006547159980982542
step: 110, loss: 0.0007810928509570658
step: 120, loss: 0.05386599153280258
step: 130, loss: 0.04074714332818985
step: 140, loss: 0.003603016957640648
step: 150, loss: 0.06635317951440811
step: 160, loss: 0.012020908296108246
step: 170, loss: 0.004334998317062855
step: 180, loss: 0.14419473707675934
step: 190, loss: 0.08214121311903
step: 200, loss: 0.016923442482948303
step: 210, loss: 0.05791740119457245
/usr/local/lib/python3.7/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=0.5811623246492985, f1=0.6049382716049383,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025067104026675224
step: 10, loss: 0.03795302286744118
step: 20, loss: 0.03156169131398201
step: 30, loss: 0.0013994361506775022
step: 40, loss: 0.006536488886922598
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 50, loss: 0.044127922505140305
step: 60, loss: 0.02988109178841114
step: 70, loss: 0.013831629417836666
step: 80, loss: 0.03571006655693054
step: 90, loss: 0.0058989860117435455
step: 100, loss: 0.007244451902806759
step: 110, loss: 0.02934345416724682
step: 120, loss: 0.08876820653676987
step: 130, loss: 0.009960015304386616
step: 140, loss: 0.012057258747518063
step: 150, loss: 0.0017973007634282112
step: 160, loss: 0.0026901846285909414
step: 170, loss: 0.01384836994111538
step: 180, loss: 0.001708770520053804
step: 190, loss: 0.01224595122039318
step: 200, loss: 0.002563005778938532
step: 210, loss: 0.009703869000077248
/usr/local/lib/python3.7/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=0.6020618556701032, f1=0.6167023554603854,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1537763476371765
step: 10, loss: 0.003674056613817811
step: 20, loss: 0.010187164880335331
step: 30, loss: 0.08767236024141312
step: 40, loss: 0.00929409172385931
step: 50, loss: 0.002654454205185175
step: 60, loss: 0.03259784355759621
step: 70, loss: 0.017160698771476746
step: 80, loss: 0.00873531773686409
step: 90, loss: 0.07227100431919098
step: 100, loss: 0.038332171738147736
step: 110, loss: 0.010493296198546886

```

```

step: 120, loss: 0.005485785659402609
step: 130, loss: 0.013367773033678532
step: 140, loss: 0.05221294239163399
step: 150, loss: 0.008460917510092258
step: 160, loss: 0.0538356676697731
step: 170, loss: 0.011115808971226215
step: 180, loss: 0.11544598639011383
step: 190, loss: 0.027789540588855743
step: 200, loss: 0.031141530722379684
step: 210, loss: 0.003419145243242383
/usr/local/lib/python3.7/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=0.6101694915254237, f1=0.5869565217391304,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044722992926836014
step: 10, loss: 0.0011472568148747087
step: 20, loss: 0.008443407714366913
step: 30, loss: 0.007295448798686266
step: 40, loss: 0.016079967841506004
step: 50, loss: 0.003967003431171179
step: 60, loss: 0.0033878714311867952
step: 70, loss: 0.1270708887691498
step: 80, loss: 0.07143152505159378
step: 90, loss: 0.0008305480587296188
step: 100, loss: 0.11170954257249832
step: 110, loss: 0.0114598935469985
step: 120, loss: 0.0016525564715266228
step: 130, loss: 0.0012524212943390012
step: 140, loss: 0.006034764461219311
step: 150, loss: 0.0035303218755871058
step: 160, loss: 0.0028589502908289433
step: 170, loss: 0.012989218346774578
step: 180, loss: 0.015447924844920635
step: 190, loss: 0.0022095399908721447
step: 200, loss: 0.2006683051586151
step: 210, loss: 0.027573762461543083
/usr/local/lib/python3.7/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=0.5957446808510638, f1=0.5970149253731344,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.0026411470025777817
step: 10, loss: 0.012885796837508678
step: 20, loss: 0.0029952186159789562
step: 30, loss: 0.016505341976881027
step: 40, loss: 0.0018286730628460646
step: 50, loss: 0.0021503979805856943
step: 60, loss: 0.005636812187731266
step: 70, loss: 0.0006380904233083129
step: 80, loss: 0.00022889654792379588
step: 90, loss: 0.007858232595026493
step: 100, loss: 0.01762227527797222
step: 110, loss: 0.001606841804459691
step: 120, loss: 0.039334263652563095
step: 130, loss: 0.13458746671676636
step: 140, loss: 0.012825156562030315
step: 150, loss: 0.05778297036886215
step: 160, loss: 0.01871071755886078
step: 170, loss: 0.003945476375520229
step: 180, loss: 0.0008398511563427746
step: 190, loss: 0.007742001209408045
step: 200, loss: 0.0005414935294538736
step: 210, loss: 0.03238924220204353
/usr/local/lib/python3.7/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=0.5936842105263158, f1=0.5739130434782609,
best_f1=0.6419753086419754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019000471802428365
step: 10, loss: 0.0003369585028849542
step: 20, loss: 0.0021067883353680372
step: 30, loss: 0.030295901000499725
step: 40, loss: 0.0012032677186653018
step: 50, loss: 0.0030139146838337183
step: 60, loss: 0.021125953644514084
step: 70, loss: 0.0006704982952214777
step: 80, loss: 0.004083804786205292
step: 90, loss: 0.039203137159347534
step: 100, loss: 0.06697995215654373
step: 110, loss: 0.0006078680162318051
step: 120, loss: 0.008364833891391754
step: 130, loss: 0.018331564962863922
step: 140, loss: 0.0011440314119681716
step: 150, loss: 0.0016042966162785888
step: 160, loss: 0.0018258177442476153
step: 170, loss: 0.004846662748605013

```

```

step: 180, loss: 0.0020231844391673803
step: 190, loss: 0.002902979264035821
step: 200, loss: 0.010898243635892868
step: 210, loss: 0.010478045791387558
/usr/local/lib/python3.7/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=0.5962732919254657, f1=0.5910064239828693,
best_f1=0.6419753086419754
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 475.58it/s]
load_f1 = 0.6481113320079521
real_f1 = 0.6387225548902197
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 262.53it/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: 0.5748521685600281
step: 10, loss: 0.37487688660621643
step: 20, loss: 0.29720446467399597
step: 30, loss: 0.39682579040527344
step: 40, loss: 0.4304918944835663
step: 50, loss: 0.2964189946651459
step: 60, loss: 0.2650407552719116
step: 70, loss: 0.29675060510635376
step: 80, loss: 0.29969140887260437
step: 90, loss: 0.2609870433807373
step: 100, loss: 0.3520505726337433
step: 110, loss: 0.5293878316879272
step: 120, loss: 0.0945090651512146
step: 130, loss: 0.11219549924135208
step: 140, loss: 0.06763942539691925
step: 150, loss: 0.18469977378845215
step: 160, loss: 0.17562930285930634
step: 170, loss: 0.32708799839019775
step: 180, loss: 0.07588095963001251
step: 190, loss: 0.2753635346889496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5529953917050692, f1=0.6082949308755761,
best_f1=0.6082949308755761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2650958299636841
step: 10, loss: 0.06543368846178055
step: 20, loss: 0.1782960295677185
step: 30, loss: 0.26542577147483826
step: 40, loss: 0.19192978739738464
step: 50, loss: 0.18320152163505554
step: 60, loss: 0.2532225549221039
step: 70, loss: 0.22741146385669708
step: 80, loss: 0.22941556572914124
step: 90, loss: 0.14478415250778198
step: 100, loss: 0.028513429686427116
step: 110, loss: 0.18424242734909058
step: 120, loss: 0.1811765730381012
step: 130, loss: 0.04680595546960831
step: 140, loss: 0.11290017515420914
step: 150, loss: 0.1747329980134964
step: 160, loss: 0.15637272596359253
step: 170, loss: 0.1667996495962143
step: 180, loss: 0.20768673717975616
step: 190, loss: 0.038575656712055206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7011494252873564, f1=0.7267605633802817,
best_f1=0.7267605633802817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04273461923003197
step: 10, loss: 0.15550465881824493
step: 20, loss: 0.061559371650218964
step: 30, loss: 0.024607524275779724
step: 40, loss: 0.08297023177146912
step: 50, loss: 0.1221911758184433
step: 60, loss: 0.03127588331699371

```

```

step: 70, loss: 0.07682093977928162
step: 80, loss: 0.14478354156017303
step: 90, loss: 0.05394483730196953
step: 100, loss: 0.035542137920856476
step: 110, loss: 0.008692593313753605
step: 120, loss: 0.0045853061601519585
step: 130, loss: 0.010006112046539783
step: 140, loss: 0.1283813714981079
step: 150, loss: 0.1666792333126068
step: 160, loss: 0.0048165637999773026
step: 170, loss: 0.1327958106994629
step: 180, loss: 0.04525178670883179
step: 190, loss: 0.10691463202238083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7319884726224785, f1=0.7486033519553073,
best_f1=0.7486033519553073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17944954335689545
step: 10, loss: 0.10731380432844162
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.1258072853088379
step: 30, loss: 0.022564100101590157
step: 40, loss: 0.04430245980620384
step: 50, loss: 0.029771601781249046
step: 60, loss: 0.139182448387146
step: 70, loss: 0.24012234807014465
step: 80, loss: 0.1182800680398941
step: 90, loss: 0.02885865606367588
step: 100, loss: 0.12413659691810608
step: 110, loss: 0.002056713216006756
step: 120, loss: 0.04267401993274689
step: 130, loss: 0.10847281664609909
step: 140, loss: 0.03269433602690697
step: 150, loss: 0.016390779986977577
step: 160, loss: 0.05688449740409851
step: 170, loss: 0.05431603267788887
step: 180, loss: 0.033221591264009476
step: 190, loss: 0.30376970767974854
/usr/local/lib/python3.7/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=0.7017543859649122, f1=0.7061855670103092,
best_f1=0.7486033519553073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05134729668498039
step: 10, loss: 0.012559439986944199
step: 20, loss: 0.061367250978946686
step: 30, loss: 0.009353847242891788
step: 40, loss: 0.11320505291223526
step: 50, loss: 0.0425712913274765
step: 60, loss: 0.13800273835659027
step: 70, loss: 0.017940182238817215
step: 80, loss: 0.006194050423800945
step: 90, loss: 0.056843455880880356
step: 100, loss: 0.007915337570011616
step: 110, loss: 0.021527450531721115
step: 120, loss: 0.011035775765776634
step: 130, loss: 0.09137826412916183
step: 140, loss: 0.19443918764591217
step: 150, loss: 0.02141113206744194
step: 160, loss: 0.06922278553247452
step: 170, loss: 0.004921182990074158
step: 180, loss: 0.06545748561620712
step: 190, loss: 0.07788655906915665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7486033519553073, f1=0.7388888888888889,
best_f1=0.7388888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010824237950146198
step: 10, loss: 0.052975356578826904
step: 20, loss: 0.0038508614525198936
step: 30, loss: 0.0003843856684397906
step: 40, loss: 0.05006714537739754
step: 50, loss: 0.004677070304751396
step: 60, loss: 0.015656691044569016
step: 70, loss: 0.12366532534360886
step: 80, loss: 0.025590483099222183
step: 90, loss: 0.013041634112596512

```

```

step: 100, loss: 0.0006083553307689726
step: 110, loss: 0.08691342175006866
step: 120, loss: 0.004115038551390171
step: 130, loss: 0.009985098615288734
step: 140, loss: 0.006054480094462633
step: 150, loss: 0.002131957793608308
step: 160, loss: 0.10262727737426758
step: 170, loss: 0.017945587635040283
step: 180, loss: 0.11650791019201279
step: 190, loss: 0.022144997492432594
/usr/local/lib/python3.7/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=0.7354497354497355, f1=0.7403598971722366,
best_f1=0.7388888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013303027488291264
step: 10, loss: 0.01837409846484661
step: 20, loss: 0.04791121929883957
step: 30, loss: 0.0023113912902772427
step: 40, loss: 0.008769119158387184
step: 50, loss: 0.01824156753718853
step: 60, loss: 0.01983848586678505
step: 70, loss: 0.03165889531373978
step: 80, loss: 0.1396162509918213
step: 90, loss: 0.15233048796653748
step: 100, loss: 0.00040083753992803395
step: 110, loss: 0.006301356013864279
step: 120, loss: 0.002064890693873167
step: 130, loss: 0.0029742103070020676
step: 140, loss: 0.01685856468975544
step: 150, loss: 0.00926282349973917
step: 160, loss: 0.02150191180408001
step: 170, loss: 0.025579784065485
step: 180, loss: 0.012115690857172012
step: 190, loss: 0.002398671815171838
/usr/local/lib/python3.7/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=0.7413333333333334, f1=0.7539267015706806,
best_f1=0.7388888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014859320595860481
step: 10, loss: 0.002940189093351364

```



```

step: 20, loss: 0.004794599022716284
step: 30, loss: 0.0007640800904482603
step: 40, loss: 0.13833820819854736
step: 50, loss: 0.16026312112808228
step: 60, loss: 0.13635897636413574
step: 70, loss: 0.008713425137102604
step: 80, loss: 0.0010229782201349735
step: 90, loss: 0.02230852283537388
step: 100, loss: 0.01409986987709999
step: 110, loss: 0.0010260010603815317
step: 120, loss: 0.00911357905715704
step: 130, loss: 0.03540800139307976
step: 140, loss: 0.010249323211610317
step: 150, loss: 0.044095177203416824
step: 160, loss: 0.024926621466875076
step: 170, loss: 0.0010819071903824806
step: 180, loss: 0.006460175849497318
step: 190, loss: 0.001601187395863235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7658402203856748, f1=0.7816711590296496,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019297258695587516
step: 10, loss: 0.008908962830901146
step: 20, loss: 0.0021454854868352413
step: 30, loss: 0.0022856947034597397
step: 40, loss: 0.0009645260870456696
step: 50, loss: 0.05505027621984482
step: 60, loss: 0.0005515672964975238
step: 70, loss: 0.0018890718929469585
step: 80, loss: 0.005094135645776987
step: 90, loss: 0.06594917178153992
step: 100, loss: 0.12385797500610352
step: 110, loss: 0.0010625669965520501
step: 120, loss: 0.012544546276330948
step: 130, loss: 0.01680082082748413
step: 140, loss: 0.01356601808220148
step: 150, loss: 0.014055243693292141
step: 160, loss: 0.0004222674760967493
step: 170, loss: 0.18307533860206604

```

```

step: 180, loss: 0.006973284762352705
step: 190, loss: 0.0031891940161585808
/usr/local/lib/python3.7/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=0.7311178247734139, f1=0.7305389221556886,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003987593576312065
step: 10, loss: 0.002297238912433386
step: 20, loss: 0.023535186424851418
step: 30, loss: 0.0006553118582814932
step: 40, loss: 0.03983166068792343
step: 50, loss: 0.001623573829419911
step: 60, loss: 0.005514620803296566
step: 70, loss: 0.005327997263520956
step: 80, loss: 0.014910672791302204
step: 90, loss: 0.07253073900938034
step: 100, loss: 0.001466405694372952
step: 110, loss: 0.008294044993817806
step: 120, loss: 0.008930621668696404
step: 130, loss: 0.026863252744078636
step: 140, loss: 0.0027883811853826046
step: 150, loss: 0.004279938526451588
step: 160, loss: 0.011900079436600208
step: 170, loss: 0.0032776903826743364
step: 180, loss: 0.0010823209304362535
step: 190, loss: 0.0013854765566065907
/usr/local/lib/python3.7/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=0.7485714285714286, f1=0.7821229050279329,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024426118470728397
step: 10, loss: 0.00129895459394902
step: 20, loss: 0.020797181874513626
step: 30, loss: 0.10322681814432144
step: 40, loss: 0.0022510376293212175
step: 50, loss: 0.020744606852531433
step: 60, loss: 0.0019037129823118448
step: 70, loss: 0.0016988675342872739
step: 80, loss: 0.0012561422772705555
step: 90, loss: 0.00040701765101403

```

```

step: 100, loss: 0.005884020589292049
step: 110, loss: 0.0012690440053120255
step: 120, loss: 0.007446755189448595
step: 130, loss: 0.0005513414507731795
step: 140, loss: 0.0029238618444651365
step: 150, loss: 0.0005674657877534628
step: 160, loss: 0.005097723565995693
step: 170, loss: 0.0003128368407487869
step: 180, loss: 0.0006609638803638518
step: 190, loss: 0.0015282955719158053
/usr/local/lib/python3.7/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=0.7567567567567567, f1=0.7780821917808218,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008041078108362854
step: 10, loss: 0.0017060736427083611
step: 20, loss: 0.00023457707720808685
step: 30, loss: 0.0008945315494202077
step: 40, loss: 0.003524672705680132
step: 50, loss: 0.003684511873871088
step: 60, loss: 0.0005048876046203077
step: 70, loss: 0.029697611927986145
step: 80, loss: 0.004142424091696739
step: 90, loss: 0.0003086104115936905
step: 100, loss: 0.0031841371674090624
step: 110, loss: 0.00038345070788636804
step: 120, loss: 0.0012877410044893622
step: 130, loss: 0.01974412240087986
step: 140, loss: 0.07120788097381592
step: 150, loss: 0.0016034294385463
step: 160, loss: 0.001338059431873262
step: 170, loss: 0.0003412103687878698
step: 180, loss: 0.0006408001645468175
step: 190, loss: 0.002943342784419656
/usr/local/lib/python3.7/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=0.7500000000000001, f1=0.7833333333333333,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003033993998542428
step: 10, loss: 0.028149448335170746

```

```

step: 20, loss: 0.0028671470936387777
step: 30, loss: 0.006977815646678209
step: 40, loss: 0.0005186983034946024
step: 50, loss: 0.150760680437088
step: 60, loss: 0.0010296839755028486
step: 70, loss: 0.003256390104070306
step: 80, loss: 0.001826481893658638
step: 90, loss: 0.006401010788977146
step: 100, loss: 0.0019301411230117083
step: 110, loss: 0.005405470263212919
step: 120, loss: 0.0008657909347675741
step: 130, loss: 0.04688909649848938
step: 140, loss: 0.0007391570834442973
step: 150, loss: 0.00026477279607206583
step: 160, loss: 0.000545013346709311
step: 170, loss: 0.12508489191532135
step: 180, loss: 0.00269303354434669
step: 190, loss: 0.02329181134700775
/usr/local/lib/python3.7/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=0.7645429362880887, f1=0.7759562841530054,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001122458023019135
step: 10, loss: 0.0324874073266983
step: 20, loss: 0.0004516587941907346
step: 30, loss: 0.0004815158899873495
step: 40, loss: 0.008893872611224651
step: 50, loss: 0.002399210352450609
step: 60, loss: 0.004001338500529528
step: 70, loss: 0.002572155324742198
step: 80, loss: 0.01528194546699524
step: 90, loss: 0.008191561326384544
step: 100, loss: 0.034939032047986984
step: 110, loss: 0.0019064489752054214
step: 120, loss: 0.002323335036635399
step: 130, loss: 0.0006177502800710499
step: 140, loss: 0.016154196113348007
step: 150, loss: 0.008881245739758015
step: 160, loss: 0.002019954612478614
step: 170, loss: 0.004089548718184233
step: 180, loss: 0.0025966253597289324
step: 190, loss: 0.0009640680509619415
/usr/local/lib/python3.7/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=0.7603305785123966, f1=0.779291553133515,
best_f1=0.7816711590296496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02592608518898487
step: 10, loss: 0.004354620818048716
step: 20, loss: 0.02433491125702858
step: 30, loss: 0.0007835496217012405
step: 40, loss: 0.0005695315194316208
step: 50, loss: 0.0020924247801303864
step: 60, loss: 0.007451374549418688
step: 70, loss: 0.02930002473294735
step: 80, loss: 0.000982025288976729
step: 90, loss: 0.0005219705053605139
step: 100, loss: 0.0009468477801419795
step: 110, loss: 0.0003993269056081772
step: 120, loss: 0.0810328796505928
step: 130, loss: 0.0955149382352829
step: 140, loss: 0.008973835967481136
step: 150, loss: 0.0004584561102092266
step: 160, loss: 0.0002905589062720537
step: 170, loss: 0.000595087418332696
step: 180, loss: 0.0009055384434759617
step: 190, loss: 0.001030450570397079
/usr/local/lib/python3.7/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=0.7601078167115903, f1=0.7780821917808218,
best_f1=0.7816711590296496
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, 223.91it/s]
load_f1 = 0.7085714285714286
real_f1 = 0.6631016042780749
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, 258.31it/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: 0.6358979940414429
step: 10, loss: 0.3804142475128174
step: 20, loss: 0.3046649694442749
step: 30, loss: 0.3934731185436249
step: 40, loss: 0.2748245298862457
step: 50, loss: 0.2683543264865875
step: 60, loss: 0.25545915961265564
step: 70, loss: 0.36577385663986206
step: 80, loss: 0.38896656036376953
step: 90, loss: 0.22536516189575195
step: 100, loss: 0.22806717455387115
step: 110, loss: 0.21211041510105133
step: 120, loss: 0.2621464133262634
step: 130, loss: 0.0862126499414444
step: 140, loss: 0.27472132444381714
step: 150, loss: 0.38891729712486267
step: 160, loss: 0.24914002418518066
step: 170, loss: 0.23842519521713257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6616161616161617, f1=0.6733167082294264,
best_f1=0.6733167082294264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1138736754655838
step: 10, loss: 0.05190000683069229
step: 20, loss: 0.12480221688747406
step: 30, loss: 0.24142640829086304
step: 40, loss: 0.1273566335439682
step: 50, loss: 0.21319644153118134
step: 60, loss: 0.10823097079992294
step: 70, loss: 0.13696718215942383
step: 80, loss: 0.056948013603687286
step: 90, loss: 0.14302237331867218
step: 100, loss: 0.1824701577425003
step: 110, loss: 0.06504753977060318
step: 120, loss: 0.12803934514522552
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.06715527921915054
step: 140, loss: 0.2671118676662445

```



```

step: 150, loss: 0.153494194149971
step: 160, loss: 0.09890249371528625
step: 170, loss: 0.11110410839319229
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7386091127098322, f1=0.7450980392156863,
best_f1=0.7450980392156863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1413058191537857
step: 10, loss: 0.07497366517782211
step: 20, loss: 0.0832546278834343
step: 30, loss: 0.18032243847846985
step: 40, loss: 0.11785591393709183
step: 50, loss: 0.2141377031803131
step: 60, loss: 0.04996515065431595
step: 70, loss: 0.08621150255203247
step: 80, loss: 0.04582102224230766
step: 90, loss: 0.2334405779838562
step: 100, loss: 0.0704159140586853
step: 110, loss: 0.15862907469272614
step: 120, loss: 0.10946521908044815
step: 130, loss: 0.11321781575679779
step: 140, loss: 0.043863337486982346
step: 150, loss: 0.0355304554104805
step: 160, loss: 0.10151653736829758
step: 170, loss: 0.08258926868438721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7653061224489796, f1=0.8058968058968059,
best_f1=0.8058968058968059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00894572027027607
step: 10, loss: 0.037053536623716354
step: 20, loss: 0.009346941486001015

```

```

step: 30, loss: 0.16489222645759583
step: 40, loss: 0.009200595319271088
step: 50, loss: 0.07923539727926254
step: 60, loss: 0.08126739412546158
step: 70, loss: 0.07119815796613693
step: 80, loss: 0.16597643494606018
step: 90, loss: 0.12189029157161713
step: 100, loss: 0.029812583699822426
step: 110, loss: 0.18060152232646942
step: 120, loss: 0.04295208305120468
step: 130, loss: 0.11322544515132904
step: 140, loss: 0.028526263311505318
step: 150, loss: 0.051589235663414
step: 160, loss: 0.47742024064064026
step: 170, loss: 0.10311673581600189
/usr/local/lib/python3.7/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=0.7540229885057472, f1=0.7755102040816327,
best_f1=0.8058968058968059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.060776639729738235
step: 10, loss: 0.028651248663663864
step: 20, loss: 0.010208523832261562
step: 30, loss: 0.014354722574353218
step: 40, loss: 0.004371557384729385
step: 50, loss: 0.018704621121287346
step: 60, loss: 0.015147336758673191
step: 70, loss: 0.18824176490306854
step: 80, loss: 0.0991934984922409
step: 90, loss: 0.05712075158953667
step: 100, loss: 0.015488861128687859
step: 110, loss: 0.16378842294216156
step: 120, loss: 0.024243544787168503
step: 130, loss: 0.1159512996673584
step: 140, loss: 0.012942780740559101
step: 150, loss: 0.04101656749844551
step: 160, loss: 0.06613294035196304
step: 170, loss: 0.020891284570097923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.772093023255814, f1=0.7671232876712328,
best_f1=0.7671232876712328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005009638611227274
step: 10, loss: 0.03316108509898186
step: 20, loss: 0.027430636808276176
step: 30, loss: 0.0043296427465975285
step: 40, loss: 0.0368974544107914
step: 50, loss: 0.21095193922519684
step: 60, loss: 0.10023988783359528
step: 70, loss: 0.11749496310949326
step: 80, loss: 0.033209867775440216
step: 90, loss: 0.032472364604473114
step: 100, loss: 0.08856700360774994
step: 110, loss: 0.0020373216830193996
step: 120, loss: 0.006394118070602417
step: 130, loss: 0.029656048864126205
step: 140, loss: 0.019671987742185593
step: 150, loss: 0.09279319643974304
step: 160, loss: 0.065308578312397
step: 170, loss: 0.015127994120121002
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7990314769975787, f1=0.7645687645687645,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007521644234657288
step: 10, loss: 0.015870239585638046
step: 20, loss: 0.00482525397092104
step: 30, loss: 0.03275414928793907
step: 40, loss: 0.014137022197246552
step: 50, loss: 0.005934370215982199
step: 60, loss: 0.007338986732065678
step: 70, loss: 0.006196944508701563
step: 80, loss: 0.048490118235349655
step: 90, loss: 0.0010006354423239827
step: 100, loss: 0.005630621686577797
step: 110, loss: 0.09419819712638855
step: 120, loss: 0.009122439660131931

```

```

step: 130, loss: 0.16234301030635834
step: 140, loss: 0.005155432969331741
step: 150, loss: 0.013772835023701191
step: 160, loss: 0.004767074249684811
step: 170, loss: 0.12491367757320404
/usr/local/lib/python3.7/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=0.7788018433179723, f1=0.7927107061503416,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018437795341014862
step: 10, loss: 0.05411003902554512
step: 20, loss: 0.004188641905784607
step: 30, loss: 0.014411705546081066
step: 40, loss: 0.0008509629406034946
step: 50, loss: 0.0037213757168501616
step: 60, loss: 0.007622744422405958
step: 70, loss: 0.015246279537677765
step: 80, loss: 0.005867003463208675
step: 90, loss: 0.024165816605091095
step: 100, loss: 0.06594154983758926
step: 110, loss: 0.12211691588163376
step: 120, loss: 0.0029590297490358353
step: 130, loss: 0.004208317492157221
step: 140, loss: 0.09588763862848282
step: 150, loss: 0.09287817776203156
step: 160, loss: 0.010723190382122993
step: 170, loss: 0.10925018042325974
/usr/local/lib/python3.7/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=0.7672634271099745, f1=0.7719298245614035,
best_f1=0.7645687645687645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01769275590777397
step: 10, loss: 0.010647936724126339
step: 20, loss: 0.007675935514271259
step: 30, loss: 0.005844771862030029
step: 40, loss: 0.0020356597378849983
step: 50, loss: 0.002025072695687413
step: 60, loss: 0.011254953220486641
step: 70, loss: 0.0032858492340892553
step: 80, loss: 0.0036437788512557745

```

```

step: 90, loss: 0.02072330005466938
step: 100, loss: 0.0401700995862484
step: 110, loss: 0.05178840085864067
step: 120, loss: 0.012771238572895527
step: 130, loss: 0.11573470383882523
step: 140, loss: 0.07880579680204391
step: 150, loss: 0.008891687728464603
step: 160, loss: 0.03027944453060627
step: 170, loss: 0.0016678563551977277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8009478672985781, f1=0.8027522935779815,
best_f1=0.8027522935779815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14516623318195343
step: 10, loss: 0.026846829801797867
step: 20, loss: 0.09176377207040787
step: 30, loss: 0.003458862891420722
step: 40, loss: 0.15373867750167847
step: 50, loss: 0.023653250187635422
step: 60, loss: 0.010723154991865158
step: 70, loss: 0.012809611856937408
step: 80, loss: 0.008451638743281364
step: 90, loss: 0.008611228317022324
step: 100, loss: 0.0005462002009153366
step: 110, loss: 0.043931953608989716
step: 120, loss: 0.008874768391251564
step: 130, loss: 0.03588220477104187
step: 140, loss: 0.06699823588132858
step: 150, loss: 0.019383950158953667
step: 160, loss: 0.041015882045030594
step: 170, loss: 0.02540004998445511
/usr/local/lib/python3.7/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=0.779746835443038, f1=0.7919799498746868,
best_f1=0.8027522935779815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020194649696350098

```

```

step: 10, loss: 0.004262535832822323
step: 20, loss: 0.031056709587574005
step: 30, loss: 0.0006208340055309236
step: 40, loss: 0.00465970765799284
step: 50, loss: 0.0004935328033752739
step: 60, loss: 0.06415057927370071
step: 70, loss: 0.02619049698114395
step: 80, loss: 0.006619841791689396
step: 90, loss: 0.002734237816184759
step: 100, loss: 0.0017957575619220734
step: 110, loss: 0.01120760478079319
step: 120, loss: 0.0036362174432724714
step: 130, loss: 0.0008764484664425254
step: 140, loss: 0.022375546395778656
step: 150, loss: 0.0028264415450394154
step: 160, loss: 0.008299865759909153
step: 170, loss: 0.017678387463092804
/usr/local/lib/python3.7/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=0.7970297029702972, f1=0.802919708029197,
best_f1=0.8027522935779815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0076223877258598804
step: 10, loss: 0.0020676234271377325
step: 20, loss: 0.013569879345595837
step: 30, loss: 0.0007191505865193903
step: 40, loss: 0.002287202747538686
step: 50, loss: 0.0073695057071745396
step: 60, loss: 0.14808598160743713
step: 70, loss: 0.08310551196336746
step: 80, loss: 0.0005109041230753064
step: 90, loss: 0.001414524856954813
step: 100, loss: 0.001856464659795165
step: 110, loss: 0.0009696583729237318
step: 120, loss: 0.0035190300550311804
step: 130, loss: 0.008285297080874443
step: 140, loss: 0.005106888245791197
step: 150, loss: 0.0037658722139894962
step: 160, loss: 0.000668843335006386
step: 170, loss: 0.0003987459640484303
/usr/local/lib/python3.7/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=0.7696078431372548, f1=0.7893462469733656,
best_f1=0.8027522935779815

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002301010536029935
step: 10, loss: 0.0005117082037031651
step: 20, loss: 0.0012367900926619768
step: 30, loss: 0.0005412223981693387
step: 40, loss: 0.0023894477635622025
step: 50, loss: 0.0005798320635221899
step: 60, loss: 0.012662220746278763
step: 70, loss: 0.0032635447569191456
step: 80, loss: 0.10046298056840897
step: 90, loss: 0.0003247610293328762
step: 100, loss: 0.008527371101081371
step: 110, loss: 0.004215881694108248
step: 120, loss: 0.014339831657707691
step: 130, loss: 0.001981359673663974
step: 140, loss: 0.0013724323362112045
step: 150, loss: 0.0022587073035538197
step: 160, loss: 0.0019471074920147657
step: 170, loss: 0.00824027694761753
/usr/local/lib/python3.7/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=0.7918781725888325, f1=0.7921760391198044,
best_f1=0.8027522935779815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012880865251645446
step: 10, loss: 0.00046141844359226525
step: 20, loss: 0.0013297925470396876
step: 30, loss: 0.0129456901922822
step: 40, loss: 0.0002870754397008568
step: 50, loss: 0.026870960369706154
step: 60, loss: 0.0007449613767676055
step: 70, loss: 0.0017353990115225315
step: 80, loss: 0.001153200282715261
step: 90, loss: 0.0008888922166079283
step: 100, loss: 0.016297198832035065
step: 110, loss: 0.005115943029522896
step: 120, loss: 0.05458776280283928
step: 130, loss: 0.011486588977277279
step: 140, loss: 0.0011826370609924197
step: 150, loss: 0.0009066889761015773
step: 160, loss: 0.002867515431717038
step: 170, loss: 0.002356958342716098
/usr/local/lib/python3.7/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=0.7888040712468193, f1=0.7990196078431373,
best_f1=0.8027522935779815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004571459721773863
step: 10, loss: 0.002555671613663435
step: 20, loss: 0.0038489431608468294
step: 30, loss: 0.0014192159287631512
step: 40, loss: 0.0115605928003788
step: 50, loss: 0.0007752456585876644
step: 60, loss: 0.03495519608259201
step: 70, loss: 0.001271246699616313
step: 80, loss: 0.007296781986951828
step: 90, loss: 0.0026544148568063974
step: 100, loss: 0.0016490871785208583
step: 110, loss: 0.00045118623529560864
step: 120, loss: 0.0004791267274413258
step: 130, loss: 0.0014314398868009448
step: 140, loss: 0.0007437804597429931
step: 150, loss: 0.0007842146442271769
step: 160, loss: 0.001799589255824685
step: 170, loss: 0.001576554961502552
/usr/local/lib/python3.7/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=0.7919799498746868, f1=0.7931873479318734,
best_f1=0.8027522935779815
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:06, 284.85it/s]
load_f1 = 0.5519591141396933
real_f1 = 0.5584642233856895
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 260.46it/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: 0.6268185973167419
step: 10, loss: 0.6247966885566711
step: 20, loss: 0.431181401014328
step: 30, loss: 0.17409704625606537
step: 40, loss: 0.2653740644454956
step: 50, loss: 0.07025144249200821
step: 60, loss: 0.048899997025728226
step: 70, loss: 0.15257588028907776
step: 80, loss: 0.044556424021720886
step: 90, loss: 0.1734917014837265
step: 100, loss: 0.0376373790204525
step: 110, loss: 0.32378003001213074
step: 120, loss: 0.013894649222493172
step: 130, loss: 0.11060351133346558
step: 140, loss: 0.010813730768859386
step: 150, loss: 0.01960681937634945
step: 160, loss: 0.015155651606619358
step: 170, loss: 0.10890638083219528
step: 180, loss: 0.021936491131782532
step: 190, loss: 0.02862461656332016
step: 200, loss: 0.11853092163801193
step: 210, loss: 0.013124142773449421
step: 220, loss: 0.006009258795529604
step: 230, loss: 0.08776308596134186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9599999999999999, f1=0.9662162162162162,
best_f1=0.9662162162162162
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03042772226035595
step: 10, loss: 0.043632835149765015
step: 20, loss: 0.1733575165271759
step: 30, loss: 0.05590850114822388
step: 40, loss: 0.11061368137598038
step: 50, loss: 0.0036073678638786077
step: 60, loss: 0.018969513475894928
step: 70, loss: 0.10132656991481781
step: 80, loss: 0.032144781202077866
step: 90, loss: 0.008044418878853321
step: 100, loss: 0.06773111969232559

```

```

step: 110, loss: 0.1768152266740799
step: 120, loss: 0.11156698316335678
step: 130, loss: 0.008513095788657665
step: 140, loss: 0.0025494606234133244
step: 150, loss: 0.033188361674547195
step: 160, loss: 0.0439625084400177
step: 170, loss: 0.0036340218503028154
step: 180, loss: 0.02226044423878193
step: 190, loss: 0.022345520555973053
step: 200, loss: 0.007894041016697884
step: 210, loss: 0.0035759785678237677
step: 220, loss: 0.02716207504272461
step: 230, loss: 0.057946767657995224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9644444444444443, f1=0.9644444444444443,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004424369428306818
step: 10, loss: 0.02306448295712471
step: 20, loss: 0.007007214706391096
step: 30, loss: 0.13712958991527557
step: 40, loss: 0.10491623729467392
step: 50, loss: 0.005799897015094757
step: 60, loss: 0.01531471498310566
step: 70, loss: 0.019669540226459503
step: 80, loss: 0.0014948571333661675
step: 90, loss: 0.060753632336854935
step: 100, loss: 0.003434245241805911
step: 110, loss: 0.001794077455997467
step: 120, loss: 0.004463411867618561
step: 130, loss: 0.0006711930618621409
step: 140, loss: 0.009841413237154484
step: 150, loss: 0.05194548889994621
step: 160, loss: 0.053370438516139984
step: 170, loss: 0.03439662605524063
step: 180, loss: 0.009777670726180077
step: 190, loss: 0.0029427772387862206
step: 200, loss: 0.038561102002859116
step: 210, loss: 0.03695598989725113
step: 220, loss: 0.001429226715117693

```

```

step: 230, loss: 0.0012540951138362288
/usr/local/lib/python3.7/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=0.9590254706533776, f1=0.970917225950783,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006275562336668372
step: 10, loss: 0.0008441006066277623
step: 20, loss: 0.0011101120617240667
step: 30, loss: 0.02866654470562935
step: 40, loss: 0.021679997444152832
step: 50, loss: 0.004882287234067917
step: 60, loss: 0.0059221540577709675
step: 70, loss: 0.0010223389836028218
step: 80, loss: 0.0014717677840963006
step: 90, loss: 0.0021472382359206676
step: 100, loss: 0.003457169746980071
step: 110, loss: 0.0012705051340162754
step: 120, loss: 0.04275009036064148
step: 130, loss: 0.004927722737193108
step: 140, loss: 0.0036811537574976683
step: 150, loss: 0.1076948493719101
step: 160, loss: 0.03826698660850525
step: 170, loss: 0.08008722960948944
step: 180, loss: 0.0009755729115568101
step: 190, loss: 0.005606337450444698
step: 200, loss: 0.014464090578258038
step: 210, loss: 0.12014798074960709
step: 220, loss: 0.0013815077254548669
step: 230, loss: 0.20374265313148499
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9697648376259798, f1=0.9697648376259798,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022277161478996277
step: 10, loss: 0.0012730089947581291
step: 20, loss: 0.01660929061472416

```

```

step: 30, loss: 0.000519410939887166
step: 40, loss: 0.0009171775891445577
step: 50, loss: 0.016108283773064613
step: 60, loss: 0.007240509148687124
step: 70, loss: 0.005857500247657299
step: 80, loss: 0.0016151488525792956
step: 90, loss: 0.005222050007432699
step: 100, loss: 0.0003830326604656875
step: 110, loss: 0.0006263061659410596
step: 120, loss: 0.00031249551102519035
step: 130, loss: 0.005030668340623379
step: 140, loss: 0.00558964628726244
step: 150, loss: 0.004065047483891249
step: 160, loss: 0.004608800634741783
step: 170, loss: 0.004051170777529478
step: 180, loss: 0.0003503016196191311
step: 190, loss: 0.1372591257095337
step: 200, loss: 0.009842142462730408
step: 210, loss: 0.000984074780717492
step: 220, loss: 0.0010006538359448314
step: 230, loss: 0.0023093107156455517
/usr/local/lib/python3.7/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=0.968609865470852, f1=0.971815107102593,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012258029542863369
step: 10, loss: 0.0016358381835743785
step: 20, loss: 0.01178350206464529
step: 30, loss: 0.0004613500786945224
step: 40, loss: 0.0005605022306554019
step: 50, loss: 0.0003767538000829518
step: 60, loss: 0.0005358420312404633
step: 70, loss: 0.00041513796895742416
step: 80, loss: 0.027978947386145592
step: 90, loss: 0.00017834786558523774
step: 100, loss: 0.05534980818629265
step: 110, loss: 0.00036079471465200186
step: 120, loss: 0.00036807722062803805
step: 130, loss: 0.0007021580822765827
step: 140, loss: 0.00037151339347474277
step: 150, loss: 0.06177107244729996
step: 160, loss: 0.022736893966794014
step: 170, loss: 0.00030009745387360454
step: 180, loss: 0.004132567439228296

```

```

step: 190, loss: 0.0007834971765987575
step: 200, loss: 0.00018944342446047813
step: 210, loss: 0.00032765051582828164
step: 220, loss: 0.00020357068569865078
step: 230, loss: 0.04698583483695984
/usr/local/lib/python3.7/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=0.9697648376259798, f1=0.9699666295884317,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000907611392904073
step: 10, loss: 0.00028120147180743515
step: 20, loss: 0.00036858898238278925
step: 30, loss: 0.0006132468697614968
step: 40, loss: 0.0004114707699045539
step: 50, loss: 0.0002476070949342102
step: 60, loss: 0.0005992381484247744
step: 70, loss: 0.04281932860612869
step: 80, loss: 0.001096912194043398
step: 90, loss: 0.0012926830677315593
step: 100, loss: 0.0007660894771106541
step: 110, loss: 0.0053980061784386635
step: 120, loss: 0.00031609940924681723
step: 130, loss: 0.0011448999866843224
step: 140, loss: 0.0037565287202596664
step: 150, loss: 0.0008793286397121847
step: 160, loss: 0.03241570666432381
step: 170, loss: 0.0005328536499291658
step: 180, loss: 0.023248106241226196
step: 190, loss: 0.0004164382116869092
step: 200, loss: 0.05820523202419281
step: 210, loss: 0.0001229563495144248
step: 220, loss: 0.03091122955083847
step: 230, loss: 0.005554980132728815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9700996677740864, f1=0.9700332963374029,
best_f1=0.9700332963374029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000301895517623052
step: 10, loss: 0.0007854574942030013
step: 20, loss: 0.0008268484962172806
step: 30, loss: 0.0010686598252505064
step: 40, loss: 0.002461395924910903
step: 50, loss: 0.0016163012478500605
step: 60, loss: 0.00042478577233850956
step: 70, loss: 0.00045695988228544593
step: 80, loss: 0.000348825502442196
step: 90, loss: 0.000696720730047673
step: 100, loss: 0.0005229242378845811
step: 110, loss: 0.029629839584231377
step: 120, loss: 0.0017694549169391394
step: 130, loss: 0.0005836461787112057
step: 140, loss: 0.0002071752678602934
step: 150, loss: 0.0001018283874145709
step: 160, loss: 0.0002999334246851504
step: 170, loss: 0.00012030345533275977
step: 180, loss: 0.00012651056749746203
step: 190, loss: 0.00010420144099043682
step: 200, loss: 0.002532476792111993
step: 210, loss: 0.00021720574295613915
step: 220, loss: 0.00010462160571478307
step: 230, loss: 0.002196486806496978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9708520179372198, f1=0.972129319955407,
best_f1=0.972129319955407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.41509031993337e-05
step: 10, loss: 0.0001187055604532361
step: 20, loss: 0.0008428939036093652
step: 30, loss: 7.571868627564982e-05
step: 40, loss: 0.02683279477059841
step: 50, loss: 0.00017273733101319522
step: 60, loss: 9.276321361539885e-05
step: 70, loss: 0.00010384216147940606
step: 80, loss: 5.145450631971471e-05
step: 90, loss: 0.00010681209823815152
step: 100, loss: 0.0001505156687926501

```



```

step: 110, loss: 3.615994137362577e-05
step: 120, loss: 5.8902427554130554e-05
step: 130, loss: 0.013781266286969185
step: 140, loss: 0.0001939079666044563
step: 150, loss: 0.008810819126665592
step: 160, loss: 0.00024715482140891254
step: 170, loss: 0.00019718968542292714
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 180, loss: 0.035100433975458145
step: 190, loss: 0.000562600907869637
step: 200, loss: 0.0001854769652709365
step: 210, loss: 0.0001293682144023478
step: 220, loss: 0.0010974983451887965
step: 230, loss: 0.005567556247115135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753363228699552, 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: 0.00012309293379075825
step: 10, loss: 0.0006566098891198635
step: 20, loss: 7.080347131704912e-05
step: 30, loss: 0.004232343751937151
step: 40, loss: 0.003956152126193047
step: 50, loss: 0.0009746179566718638
step: 60, loss: 0.0002583756868261844
step: 70, loss: 0.0007707460899837315
step: 80, loss: 9.18918231036514e-05
step: 90, loss: 0.1488073319196701
step: 100, loss: 0.0006029705982655287
step: 110, loss: 0.0003258445649407804
step: 120, loss: 0.0006634207675233483
step: 130, loss: 0.020126106217503548
step: 140, loss: 0.03333708271384239
step: 150, loss: 0.019209640100598335
step: 160, loss: 0.00016083660011645406
step: 170, loss: 0.00014680753520224243
step: 180, loss: 0.0008461043471470475
step: 190, loss: 0.009220442734658718
step: 200, loss: 5.248318120720796e-05
step: 210, loss: 0.00033898805850185454

```

```

step: 220, loss: 0.0001217383032781072
step: 230, loss: 0.00017717447190079838
/usr/local/lib/python3.7/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=0.9718785151856018, 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: 0.001082547940313816
step: 10, loss: 0.0007542225066572428
step: 20, loss: 0.0006373166688717902
step: 30, loss: 0.0032065820414572954
step: 40, loss: 0.00023755579604767263
step: 50, loss: 0.0003005600592587143
step: 60, loss: 0.0005826543783769011
step: 70, loss: 0.0002473366039339453
step: 80, loss: 6.826022581662983e-05
step: 90, loss: 0.0001704392780084163
step: 100, loss: 6.0621227021329105e-05
step: 110, loss: 0.013467181473970413
step: 120, loss: 0.00011571295181056485
step: 130, loss: 4.7608460590709e-05
step: 140, loss: 0.00017506699077785015
step: 150, loss: 0.033431556075811386
step: 160, loss: 0.00011749450641218573
step: 170, loss: 0.009234398603439331
step: 180, loss: 0.0025361478328704834
step: 190, loss: 5.618896466330625e-05
step: 200, loss: 0.00101162726059556
step: 210, loss: 5.147333286004141e-05
step: 220, loss: 9.098344162339345e-05
step: 230, loss: 9.956162830349058e-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=0.9711111111111111, f1=0.9711751662971175,
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: 0.0016272625653073192
step: 10, loss: 0.00023416437034029514
step: 20, loss: 8.574886305723339e-05
step: 30, loss: 0.0021005223970860243
step: 40, loss: 0.00026614920352585614
step: 50, loss: 0.0015271123265847564

```

```

step: 60, loss: 0.0009461512672714889
step: 70, loss: 8.716042066225782e-05
step: 80, loss: 0.0006352121708914638
step: 90, loss: 6.405540625564754e-05
step: 100, loss: 0.0001206728775287047
step: 110, loss: 0.01903690956532955
step: 120, loss: 9.165136725641787e-05
step: 130, loss: 0.0003156965831294656
step: 140, loss: 0.0031028420198708773
step: 150, loss: 0.001065470976755023
step: 160, loss: 5.886482176720165e-05
step: 170, loss: 0.0001407195086358115
step: 180, loss: 0.03380024805665016
step: 190, loss: 0.010658049955964088
step: 200, loss: 3.465500776655972e-05
step: 210, loss: 6.48590867058374e-05
step: 220, loss: 5.702258931705728e-05
step: 230, loss: 0.03520560637116432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9763779527559054, f1=0.9763779527559054,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000746986479498446
step: 10, loss: 0.0002378136559855193
step: 20, loss: 0.00016260403208434582
step: 30, loss: 0.0002799673820845783
step: 40, loss: 0.00017760424816515297
step: 50, loss: 0.000406405859393999
step: 60, loss: 0.00013225697330199182
step: 70, loss: 0.0001730285002849996
step: 80, loss: 5.890129614272155e-05
step: 90, loss: 0.0017442188691347837
step: 100, loss: 4.487248952500522e-05
step: 110, loss: 0.016238968819379807
step: 120, loss: 0.022636588662862778
step: 130, loss: 4.498849739320576e-05
step: 140, loss: 6.942627805983648e-05
step: 150, loss: 5.757591497967951e-05
step: 160, loss: 0.00014203315367922187
step: 170, loss: 0.0001559929078212008

```

```

step: 180, loss: 4.7712404921185225e-05
step: 190, loss: 0.00021155660215299577
step: 200, loss: 5.216161662247032e-05
step: 210, loss: 2.5178753276122734e-05
step: 220, loss: 0.0009613155853003263
step: 230, loss: 3.14851786242798e-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=0.9762174405436014, f1=0.976271186440678,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011008769797626883
step: 10, loss: 0.0003261104866396636
step: 20, loss: 0.00015060327132232487
step: 30, loss: 5.9371352108428255e-05
step: 40, loss: 0.019649265334010124
step: 50, loss: 6.536687578773126e-05
step: 60, loss: 6.015353574184701e-05
step: 70, loss: 0.003872416913509369
step: 80, loss: 0.0002312414871994406
step: 90, loss: 6.751132605131716e-05
step: 100, loss: 5.4874777561053634e-05
step: 110, loss: 0.0016193098854273558
step: 120, loss: 3.8137073715915903e-05
step: 130, loss: 0.00018042774172499776
step: 140, loss: 0.0025702188722789288
step: 150, loss: 5.646973295370117e-05
step: 160, loss: 0.010772457346320152
step: 170, loss: 4.6063480112934485e-05
step: 180, loss: 6.530081009259447e-05
step: 190, loss: 0.00013166297867428511
step: 200, loss: 5.100499402033165e-05
step: 210, loss: 0.06368536502122879
step: 220, loss: 0.00019043976499233395
step: 230, loss: 0.034433912485837936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9751693002257337, f1=0.9775280898876404,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010028282849816605
step: 10, loss: 4.347638241597451e-05

```

```

step: 20, loss: 8.474405331071466e-05
step: 30, loss: 0.00011707057274179533
step: 40, loss: 0.0001518966892035678
step: 50, loss: 0.02318352460861206
step: 60, loss: 8.581506699556485e-05
step: 70, loss: 0.00019724649609997869
step: 80, loss: 0.0006677304045297205
step: 90, loss: 0.0006172873545438051
step: 100, loss: 0.0005846224376000464
step: 110, loss: 0.0005989767960272729
step: 120, loss: 0.0032325678039342165
step: 130, loss: 0.004524335265159607
step: 140, loss: 4.7724814066896215e-05
step: 150, loss: 0.00016644033894408494
step: 160, loss: 3.141425986541435e-05
step: 170, loss: 5.056156078353524e-05
step: 180, loss: 0.00011269754759268835
step: 190, loss: 0.0002570237556938082
step: 200, loss: 4.9553618737263605e-05
step: 210, loss: 0.00011251454998273402
step: 220, loss: 0.011925182305276394
step: 230, loss: 3.706071947817691e-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=0.976271186440678, f1=0.9775280898876404,
best_f1=0.9763779527559054
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 216.45it/s]
load_f1 = 0.9744160177975528
real_f1 = 0.9700996677740864
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, 256.98it/s]

```

4.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

```

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

```

Downloading: 100% 433/433 [00:00<00:00, 372kB/s]
 Downloading: 100% 232k/232k [00:00<00:00, 340kB/s]
 Downloading: 100% 440M/440M [00:23<00:00, 18.8MB/s]
 Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

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

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

step: 0, loss: 0.6263441443443298

step: 10, loss: 0.5422959923744202
step: 20, loss: 0.538320779800415
step: 30, loss: 0.28129613399505615
step: 40, loss: 0.18666069209575653
step: 50, loss: 0.05580400675535202
step: 60, loss: 0.08117463439702988
step: 70, loss: 0.21542014181613922
step: 80, loss: 0.05957085266709328
step: 90, loss: 0.2760779559612274
step: 100, loss: 0.06588171422481537
step: 110, loss: 0.08823135495185852
step: 120, loss: 0.11532507836818695
step: 130, loss: 0.10260006040334702
step: 140, loss: 0.017330430448055267
step: 150, loss: 0.04468473792076111
step: 160, loss: 0.05760886147618294
step: 170, loss: 0.24745383858680725
step: 180, loss: 0.07650333642959595
step: 190, loss: 0.03370704874396324
step: 200, loss: 0.11344972252845764
step: 210, loss: 0.08719383180141449
step: 220, loss: 0.24003532528877258
step: 230, loss: 0.17931783199310303
step: 240, loss: 0.04055439308285713
step: 250, loss: 0.038576625287532806
step: 260, loss: 0.17586012184619904
step: 270, loss: 0.03328678384423256
step: 280, loss: 0.09782367199659348
step: 290, loss: 0.10200545191764832
step: 300, loss: 0.032059408724308014
step: 310, loss: 0.18019340932369232
step: 320, loss: 0.14011983573436737
step: 330, loss: 0.05250430852174759
step: 340, loss: 0.06215352192521095
step: 350, loss: 0.01249674241989851
step: 360, loss: 0.19167301058769226
step: 370, loss: 0.115422323346138
step: 380, loss: 0.048335183411836624
step: 390, loss: 0.12309517711400986
step: 400, loss: 0.2062489092350006
step: 410, loss: 0.04180289804935455
step: 420, loss: 0.027076300233602524
step: 430, loss: 0.12122359871864319
step: 440, loss: 0.04461876302957535
step: 450, loss: 0.02117585577070713
step: 460, loss: 0.005673619918525219
step: 470, loss: 0.07085166126489639
step: 480, loss: 0.10046060383319855


```

step: 490, loss: 0.0627736747264862
step: 500, loss: 0.13356377184391022
step: 510, loss: 0.04027629643678665
step: 520, loss: 0.26647230982780457
step: 530, loss: 0.007560840807855129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9231494578029231, f1=0.9227166276346604,
best_f1=0.9227166276346604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0741306021809578
step: 10, loss: 0.09889821708202362
step: 20, loss: 0.049205850809812546
step: 30, loss: 0.21569184958934784
step: 40, loss: 0.05649994686245918
step: 50, loss: 0.09137087315320969
step: 60, loss: 0.01186579093337059
step: 70, loss: 0.1777612566947937
step: 80, loss: 0.10539966821670532
step: 90, loss: 0.03601565584540367
step: 100, loss: 0.08990199863910675
step: 110, loss: 0.12431672215461731
step: 120, loss: 0.1028200089931488
step: 130, loss: 0.0400862880051136
step: 140, loss: 0.1871359795331955
step: 150, loss: 0.07888779044151306
step: 160, loss: 0.038981884717941284
step: 170, loss: 0.24662235379219055
step: 180, loss: 0.020145995542407036
step: 190, loss: 0.08211145550012589
step: 200, loss: 0.006461150478571653
step: 210, loss: 0.0453551821410656
step: 220, loss: 0.06555742025375366
step: 230, loss: 0.07325336337089539
step: 240, loss: 0.029503513127565384
step: 250, loss: 0.04258858785033226
step: 260, loss: 0.051397740840911865
step: 270, loss: 0.14577752351760864
step: 280, loss: 0.03237168490886688
step: 290, loss: 0.0446699894964695
step: 300, loss: 0.11270115524530411

```

```

step: 310, loss: 0.008886703290045261
step: 320, loss: 0.15687011182308197
step: 330, loss: 0.059792399406433105
step: 340, loss: 0.030367476865649223
step: 350, loss: 0.006892746780067682
step: 360, loss: 0.11750373244285583
step: 370, loss: 0.14690572023391724
step: 380, loss: 0.04403352737426758
step: 390, loss: 0.0839281752705574
step: 400, loss: 0.07118800282478333
step: 410, loss: 0.017694644629955292
step: 420, loss: 0.05997280403971672
step: 430, loss: 0.04019733518362045
step: 440, loss: 0.09397319704294205
step: 450, loss: 0.03954606503248215
step: 460, loss: 0.02838226407766342
step: 470, loss: 0.07848024368286133
step: 480, loss: 0.2093539535999298
step: 490, loss: 0.02477206476032734
step: 500, loss: 0.26993781328201294
step: 510, loss: 0.024231867864727974
step: 520, loss: 0.08033664524555206
step: 530, loss: 0.010186073370277882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9299719887955181, f1=0.9188935771214253,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13743431866168976
step: 10, loss: 0.01770239882171154
step: 20, loss: 0.030158163979649544
step: 30, loss: 0.05812908336520195
step: 40, loss: 0.004175954964011908
step: 50, loss: 0.17334023118019104
step: 60, loss: 0.01351814717054367
step: 70, loss: 0.006282889749854803
step: 80, loss: 0.004224184900522232
step: 90, loss: 0.0065091815777122974
step: 100, loss: 0.0680222362279892
step: 110, loss: 0.0024481010623276234
step: 120, loss: 0.002165780169889331

```

```

step: 130, loss: 0.0020156907849013805
step: 140, loss: 0.02632877230644226
step: 150, loss: 0.14535880088806152
step: 160, loss: 0.005936040543019772
step: 170, loss: 0.06646917015314102
step: 180, loss: 0.04056720808148384
step: 190, loss: 0.009109395556151867
step: 200, loss: 0.0660613402724266
step: 210, loss: 0.065157949924469
step: 220, loss: 0.03729584068059921
step: 230, loss: 0.020149964839220047
step: 240, loss: 0.02297905460000038
step: 250, loss: 0.034921448677778244
step: 260, loss: 0.018696963787078857
step: 270, loss: 0.03055976890027523
step: 280, loss: 0.06840650737285614
step: 290, loss: 0.005438537336885929
step: 300, loss: 0.11902067065238953
step: 310, loss: 0.005298024974763393
step: 320, loss: 0.018929840996861458
step: 330, loss: 0.04057563096284866
step: 340, loss: 0.03281598910689354
step: 350, loss: 0.004518305882811546
step: 360, loss: 0.007606314029544592
step: 370, loss: 0.021685125306248665
step: 380, loss: 0.017247198149561882
step: 390, loss: 0.019718112424016
step: 400, loss: 0.0265974048525095
step: 410, loss: 0.01043847668915987
step: 420, loss: 0.1630028635263443
step: 430, loss: 0.02904169261455536
step: 440, loss: 0.07193893939256668
step: 450, loss: 0.07647436857223511
step: 460, loss: 0.01765414886176586
step: 470, loss: 0.03838871046900749
step: 480, loss: 0.006192476022988558
step: 490, loss: 0.0017909867456182837
step: 500, loss: 0.003522126702591777
step: 510, loss: 0.009406518191099167
step: 520, loss: 0.1844681352376938
step: 530, loss: 0.04791044816374779
/usr/local/lib/python3.7/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=0.9275766016713091, f1=0.9199255121042831,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026366176083683968
step: 10, loss: 0.024726320058107376
step: 20, loss: 0.003292656969279051
step: 30, loss: 0.0021639468614012003
step: 40, loss: 0.04543953016400337
step: 50, loss: 0.007482662331312895
step: 60, loss: 0.010970402508974075
step: 70, loss: 0.014376931823790073
step: 80, loss: 0.009107978083193302
step: 90, loss: 0.05482754856348038
step: 100, loss: 0.022996000945568085
step: 110, loss: 0.0010195377981290221
step: 120, loss: 0.0009471333469264209
step: 130, loss: 0.0018285212572664022
step: 140, loss: 0.007301033474504948
step: 150, loss: 0.0032989648170769215
step: 160, loss: 0.07408155500888824
step: 170, loss: 0.012756699696183205
step: 180, loss: 0.015437295660376549
step: 190, loss: 0.03448645398020744
step: 200, loss: 0.013406487181782722
step: 210, loss: 0.007601963356137276
step: 220, loss: 0.059350114315748215
step: 230, loss: 0.10126177221536636
step: 240, loss: 0.02409338392317295
step: 250, loss: 0.02651871368288994
step: 260, loss: 0.035530004650354385
step: 270, loss: 0.023804837837815285
step: 280, loss: 0.041852861642837524
step: 290, loss: 0.04186892509460449
step: 300, loss: 0.00033584743505343795
step: 310, loss: 0.05305545777082443
step: 320, loss: 0.15026044845581055
step: 330, loss: 0.145939439535141
step: 340, loss: 0.025162385776638985
step: 350, loss: 0.030813319608569145
step: 360, loss: 0.02777247317135334
step: 370, loss: 0.005911823362112045
step: 380, loss: 0.02194920741021633
step: 390, loss: 0.03662055358290672
step: 400, loss: 0.0280385110527277
step: 410, loss: 0.006120931822806597
step: 420, loss: 0.011810424737632275
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 0.15639127790927887
step: 440, loss: 0.078633151948452
step: 450, loss: 0.0232026819139719
```

```

step: 460, loss: 0.005140499211847782
step: 470, loss: 0.027711516246199608
step: 480, loss: 0.06583379209041595
step: 490, loss: 0.007621748372912407
step: 500, loss: 0.004234509542584419
step: 510, loss: 0.07544154673814774
step: 520, loss: 0.06249309703707695
step: 530, loss: 0.06611932069063187
/usr/local/lib/python3.7/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=0.925207756232687, f1=0.9162011173184358,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1296815276145935
step: 10, loss: 0.03197420760989189
step: 20, loss: 0.0029952405020594597
step: 30, loss: 0.003895822213962674
step: 40, loss: 0.009630519896745682
step: 50, loss: 0.03153996542096138
step: 60, loss: 0.14783278107643127
step: 70, loss: 0.020845919847488403
step: 80, loss: 0.0037032747641205788
step: 90, loss: 0.00808038655668497
step: 100, loss: 0.0046204919926822186
step: 110, loss: 0.004851222038269043
step: 120, loss: 0.004586380906403065
step: 130, loss: 0.0004957496421411633
step: 140, loss: 0.003411268349736929
step: 150, loss: 0.005656341556459665
step: 160, loss: 0.0005953649524599314
step: 170, loss: 0.009735982865095139
step: 180, loss: 0.007014965172857046
step: 190, loss: 0.0027323788963258266
step: 200, loss: 0.0026415123138576746
step: 210, loss: 0.0009820784907788038
step: 220, loss: 0.0003298665687907487
step: 230, loss: 0.008007308468222618
step: 240, loss: 0.05776214972138405
step: 250, loss: 0.08681215345859528
step: 260, loss: 0.007012795656919479
step: 270, loss: 0.002282913541421294
step: 280, loss: 0.0032536920625716448
step: 290, loss: 0.1404784470796585
step: 300, loss: 0.002164955250918865
step: 310, loss: 0.009943959303200245

```

```

step: 320, loss: 0.1695280373096466
step: 330, loss: 0.2209816575050354
step: 340, loss: 0.00922276172786951
step: 350, loss: 0.007139923050999641
step: 360, loss: 0.012188193388283253
step: 370, loss: 0.05542575567960739
step: 380, loss: 0.0010875500738620758
step: 390, loss: 0.0004985832492820919
step: 400, loss: 0.001076212851330638
step: 410, loss: 0.0034615658223628998
step: 420, loss: 0.0010999965015798807
step: 430, loss: 0.008532071486115456
step: 440, loss: 0.01628212071955204
step: 450, loss: 0.006901222746819258
step: 460, loss: 0.12064046412706375
step: 470, loss: 0.043055661022663116
step: 480, loss: 0.03439009562134743
step: 490, loss: 0.003993666730821133
step: 500, loss: 0.014692233875393867
step: 510, loss: 0.1322866827249527
step: 520, loss: 0.011275452561676502
step: 530, loss: 0.016567591577768326
/usr/local/lib/python3.7/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=0.9239981575310916, f1=0.9218821379625399,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063691288232803345
step: 10, loss: 0.00019652696209959686
step: 20, loss: 0.029393384233117104
step: 30, loss: 0.0011084831785410643
step: 40, loss: 0.02714735083281994
step: 50, loss: 0.03097156062722206
step: 60, loss: 0.005932871717959642
step: 70, loss: 0.00661913026124239
step: 80, loss: 0.05607961490750313
step: 90, loss: 0.014363656751811504
step: 100, loss: 0.0066025578416883945
step: 110, loss: 0.020142776891589165
step: 120, loss: 0.006160982884466648
step: 130, loss: 0.0006805055891163647
step: 140, loss: 0.005771709606051445
step: 150, loss: 0.0008612805395387113
step: 160, loss: 0.002088561886921525
step: 170, loss: 0.0017223403556272388

```

```

step: 180, loss: 0.0016568029532209039
step: 190, loss: 0.06934888660907745
step: 200, loss: 0.007234266959130764
step: 210, loss: 0.011068717576563358
step: 220, loss: 0.016518373042345047
step: 230, loss: 0.0015606287634000182
step: 240, loss: 0.04019057750701904
step: 250, loss: 0.20633651316165924
step: 260, loss: 0.014120636507868767
step: 270, loss: 0.10021437704563141
step: 280, loss: 0.021427780389785767
step: 290, loss: 0.0007402130286209285
step: 300, loss: 0.0002792020095512271
step: 310, loss: 0.006749187596142292
step: 320, loss: 0.0010196005459874868
step: 330, loss: 0.00033122641616500914
step: 340, loss: 0.027045907452702522
step: 350, loss: 0.027519894763827324
step: 360, loss: 0.014506897889077663
step: 370, loss: 0.003855054033920169
step: 380, loss: 0.0021626525558531284
step: 390, loss: 0.0019924850203096867
step: 400, loss: 0.0007444850052706897
step: 410, loss: 0.006940852850675583
step: 420, loss: 0.0033410806208848953
step: 430, loss: 0.0005822185194119811
step: 440, loss: 0.0007540853694081306
step: 450, loss: 0.00013352853420656174
step: 460, loss: 0.0002689276298042387
step: 470, loss: 0.08072850108146667
step: 480, loss: 0.0018351171165704727
step: 490, loss: 0.006130819208920002
step: 500, loss: 0.002241765847429633
step: 510, loss: 0.00488983653485775
step: 520, loss: 0.009333032183349133
step: 530, loss: 0.010938185267150402
/usr/local/lib/python3.7/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=0.9172644667623147, f1=0.9155597722960152,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004841187037527561
step: 10, loss: 0.0003756880760192871
step: 20, loss: 0.005426596384495497
step: 30, loss: 0.047264933586120605

```

step: 40, loss: 0.01193028874695301
step: 50, loss: 0.001998870400711894
step: 60, loss: 0.0026355083100497723
step: 70, loss: 0.06317660212516785
step: 80, loss: 0.09725791960954666
step: 90, loss: 0.0005669857491739094
step: 100, loss: 0.00024932328960858285
step: 110, loss: 0.0008421456441283226
step: 120, loss: 0.008377330377697945
step: 130, loss: 0.002331276424229145
step: 140, loss: 0.0004238022374920547
step: 150, loss: 0.026409268379211426
step: 160, loss: 0.00013214029604569077
step: 170, loss: 0.0003122409398201853
step: 180, loss: 0.0026237976271659136
step: 190, loss: 0.06748253852128983
step: 200, loss: 0.00649880850687623
step: 210, loss: 0.05998358502984047
step: 220, loss: 0.0043152146972715855
step: 230, loss: 0.0642395094037056
step: 240, loss: 0.00124158279504627
step: 250, loss: 0.0023616510443389416
step: 260, loss: 0.0005122327129356563
step: 270, loss: 0.0013246553717181087
step: 280, loss: 0.0013361196033656597
step: 290, loss: 0.00013178290100768209
step: 300, loss: 0.0033520134165883064
step: 310, loss: 0.00013529454008676112
step: 320, loss: 0.0010549018625169992
step: 330, loss: 0.0009608073742128909
step: 340, loss: 0.001096151303499937
step: 350, loss: 0.00031961678178049624
step: 360, loss: 0.002277084393426776
step: 370, loss: 0.0005786217516288161
step: 380, loss: 0.17880846560001373
step: 390, loss: 0.0007124514086171985
step: 400, loss: 0.0023822966031730175
step: 410, loss: 0.0018701753579080105
step: 420, loss: 0.0005827357526868582
step: 430, loss: 0.011550122871994972
step: 440, loss: 0.0010663775028660893
step: 450, loss: 0.012017160654067993
step: 460, loss: 0.0012723710387945175
step: 470, loss: 0.0008866938296705484
step: 480, loss: 0.19309945404529572
step: 490, loss: 0.007274719886481762
step: 500, loss: 0.0015072921523824334
step: 510, loss: 0.005241123028099537


```

step: 520, loss: 0.001692006946541369
step: 530, loss: 0.0003959685273002833
/usr/local/lib/python3.7/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=0.9222118088097471, f1=0.9210649229332087,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005318632815033197
step: 10, loss: 0.017370056360960007
step: 20, loss: 0.022777236998081207
step: 30, loss: 0.0021280485671013594
step: 40, loss: 0.0005411102320067585
step: 50, loss: 0.0036571354139596224
step: 60, loss: 0.001617704750970006
step: 70, loss: 0.0007573111797682941
step: 80, loss: 0.025031523779034615
step: 90, loss: 0.0006895968108437955
step: 100, loss: 0.00024584136554040015
step: 110, loss: 0.001432926394045353
step: 120, loss: 0.014634670689702034
step: 130, loss: 0.0006284479168243706
step: 140, loss: 0.0006703464314341545
step: 150, loss: 0.001215828931890428
step: 160, loss: 0.0016642767004668713
step: 170, loss: 0.014393370598554611
step: 180, loss: 0.00038603608845733106
step: 190, loss: 0.0029081341344863176
step: 200, loss: 0.0017589522758498788
step: 210, loss: 0.0027652315329760313
step: 220, loss: 0.027667183429002762
step: 230, loss: 0.0029490788001567125
step: 240, loss: 0.0001783768821042031
step: 250, loss: 0.02071077935397625
step: 260, loss: 0.0012211843859404325
step: 270, loss: 0.00037323805736377835
step: 280, loss: 0.006989312823861837
step: 290, loss: 0.0003958796733058989
step: 300, loss: 0.0014192434027791023
step: 310, loss: 0.0006760176038369536
step: 320, loss: 0.050158243626356125
step: 330, loss: 0.009695669636130333
step: 340, loss: 0.0006764255813322961
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 350, loss: 0.01430059876292944
step: 360, loss: 0.003760232590138912

```

```

step: 370, loss: 0.008350121788680553
step: 380, loss: 0.03538047894835472
step: 390, loss: 0.14317041635513306
step: 400, loss: 0.0008646230562590063
step: 410, loss: 0.000493479601573199
step: 420, loss: 0.001505960477516055
step: 430, loss: 0.02895239181816578
step: 440, loss: 0.019907843321561813
step: 450, loss: 0.001104547642171383
step: 460, loss: 0.00023008558491710573
step: 470, loss: 0.0011073397472500801
step: 480, loss: 0.004999987781047821
step: 490, loss: 0.023940976709127426
step: 500, loss: 0.0018561682663857937
step: 510, loss: 0.00022111025464255363
step: 520, loss: 0.001934269443154335
step: 530, loss: 0.0002829242730513215
/usr/local/lib/python3.7/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=0.9252506836827713, f1=0.9137614678899082,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005861541139893234
step: 10, loss: 0.013211864978075027
step: 20, loss: 0.0032082144170999527
step: 30, loss: 0.0009906067280098796
step: 40, loss: 0.01636212132871151
step: 50, loss: 0.00012475071707740426
step: 60, loss: 0.00010707540059229359
step: 70, loss: 0.024846483021974564
step: 80, loss: 0.0019516923930495977
step: 90, loss: 0.00011088103929068893
step: 100, loss: 0.00012236267502885312
step: 110, loss: 0.00033133153920061886
step: 120, loss: 0.0007630155887454748
step: 130, loss: 0.002594588091596961
step: 140, loss: 0.0017904838314279914
step: 150, loss: 0.012734124436974525
step: 160, loss: 0.0015585578512400389
step: 170, loss: 0.010204257443547249
step: 180, loss: 8.335476013598964e-05
step: 190, loss: 0.0005703337374143302
step: 200, loss: 0.00020308923558332026
step: 210, loss: 0.0015582535415887833
step: 220, loss: 0.06984526664018631

```

```

step: 230, loss: 0.00015542951587121934
step: 240, loss: 0.0003639111528173089
step: 250, loss: 0.00948353111743927
step: 260, loss: 0.09115090221166611
step: 270, loss: 0.0001681304129306227
step: 280, loss: 0.0002498587127774954
step: 290, loss: 0.12917639315128326
step: 300, loss: 0.002674719551578164
step: 310, loss: 0.010633496567606926
step: 320, loss: 0.06189614534378052
step: 330, loss: 0.0006219984497874975
step: 340, loss: 0.037941157817840576
step: 350, loss: 0.00949420128017664
step: 360, loss: 0.009729929268360138
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 370, loss: 0.002213601255789399
step: 380, loss: 0.013033438473939896
step: 390, loss: 0.0005703230272047222
step: 400, loss: 0.03762536123394966
step: 410, loss: 0.00010898045729845762
step: 420, loss: 0.00013216760999057442
step: 430, loss: 9.297537326347083e-05
step: 440, loss: 0.0021905102767050266
step: 450, loss: 0.00017416744958609343
step: 460, loss: 0.0005382007802836597
step: 470, loss: 0.0001227996835950762
step: 480, loss: 0.0004956107004545629
step: 490, loss: 0.01934344507753849
step: 500, loss: 0.0013279395643621683
step: 510, loss: 0.01455722376704216
step: 520, loss: 0.00026598613476380706
step: 530, loss: 4.731606532004662e-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=0.9222222222222223, f1=0.9177891314444496,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00723790330812335
step: 10, loss: 0.00014433912292588502
step: 20, loss: 9.483093162998557e-05
step: 30, loss: 0.0004406372900120914
step: 40, loss: 0.00019988366693723947
step: 50, loss: 0.0006521041505038738
step: 60, loss: 0.00015310382877942175
step: 70, loss: 8.264046482509002e-05

```

```
step: 80, loss: 0.0002667394292075187
step: 90, loss: 0.02997969463467598
step: 100, loss: 0.009776776656508446
step: 110, loss: 6.857756670797244e-05
step: 120, loss: 0.0004380159662105143
step: 130, loss: 0.00010730099165812135
step: 140, loss: 0.006344191264361143
step: 150, loss: 7.477691542590037e-05
step: 160, loss: 3.815909440163523e-05
step: 170, loss: 0.0003318810777273029
step: 180, loss: 6.557336018886417e-05
step: 190, loss: 0.0004798550216946751
step: 200, loss: 0.01101915817707777
step: 210, loss: 0.0030316519550979137
step: 220, loss: 0.006839345674961805
step: 230, loss: 0.0008719592005945742
step: 240, loss: 0.0018691099248826504
step: 250, loss: 0.13305984437465668
step: 260, loss: 0.0013670481275767088
step: 270, loss: 0.00113234284799546
step: 280, loss: 0.00023450089793186635
step: 290, loss: 0.005090638063848019
step: 300, loss: 7.586859283037484e-05
step: 310, loss: 0.00046783447032794356
step: 320, loss: 0.011276337318122387
step: 330, loss: 0.024186620488762856
step: 340, loss: 0.000538929714821279
step: 350, loss: 0.00015781256661284715
step: 360, loss: 0.00013305575703270733
step: 370, loss: 0.0014496354851871729
step: 380, loss: 0.01204930804669857
step: 390, loss: 0.00036337695200927556
step: 400, loss: 0.0011833587195724249
step: 410, loss: 0.001779036014340818
step: 420, loss: 0.0007949534337967634
step: 430, loss: 0.0020867050625383854
step: 440, loss: 0.010790344327688217
step: 450, loss: 0.0011327924439683557
step: 460, loss: 0.0006373940268531442
step: 470, loss: 0.0006492443499155343
step: 480, loss: 0.000385663122870028
step: 490, loss: 0.001918478519655764
step: 500, loss: 0.0061456928960978985
step: 510, loss: 0.00013074224989395589
step: 520, loss: 0.0001753668620949611
step: 530, loss: 0.00044851849088445306
/usr/local/lib/python3.7/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=0.9291044776119404, f1=0.9279999999999999,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014537631068378687
step: 10, loss: 0.22688867151737213
step: 20, loss: 0.003011011052876711
step: 30, loss: 0.01587207242846489
step: 40, loss: 0.00040300432010553777
step: 50, loss: 0.004064689856022596
step: 60, loss: 0.0050233579240739346
step: 70, loss: 0.0020459885708987713
step: 80, loss: 0.00013651378685608506
step: 90, loss: 0.005565132014453411
step: 100, loss: 0.003401447320356965
step: 110, loss: 0.0006355591467581689
step: 120, loss: 0.0005101982387714088
step: 130, loss: 6.565609510289505e-05
step: 140, loss: 0.12443101406097412
step: 150, loss: 2.7052472432842478e-05
step: 160, loss: 0.005066373385488987
step: 170, loss: 0.0003255298943258822
step: 180, loss: 0.0003073572588618845
step: 190, loss: 0.0001260759454453364
step: 200, loss: 0.002057983772829175
step: 210, loss: 0.0011973886284977198
step: 220, loss: 0.0016976551851257682
step: 230, loss: 5.006661740480922e-05
step: 240, loss: 0.000447752681793645
step: 250, loss: 3.396193642402068e-05
step: 260, loss: 5.7342967920703813e-05
step: 270, loss: 0.00026335075381211936
step: 280, loss: 0.00014905950229149312
step: 290, loss: 0.00018272604211233556
step: 300, loss: 0.00016024208161979914
step: 310, loss: 3.871173248626292e-05
step: 320, loss: 4.207608799333684e-05
step: 330, loss: 0.0005184632609598339
step: 340, loss: 0.0005878841038793325
step: 350, loss: 7.513495074817911e-05
step: 360, loss: 9.319264790974557e-05
step: 370, loss: 0.003980282228440046
step: 380, loss: 0.014624882489442825
step: 390, loss: 6.28065608907491e-05
step: 400, loss: 0.0018270923756062984
step: 410, loss: 4.556407293421216e-05
```

```
step: 420, loss: 0.00020385657262522727
step: 430, loss: 1.8719325453275815e-05
step: 440, loss: 0.00719557236880064
step: 450, loss: 5.073681313660927e-05
step: 460, loss: 0.015662625432014465
step: 470, loss: 2.4012746507651173e-05
step: 480, loss: 0.0014506785664707422
step: 490, loss: 0.00015375937800854445
step: 500, loss: 0.00015883444575592875
step: 510, loss: 0.07460447400808334
step: 520, loss: 5.252924165688455e-05
step: 530, loss: 0.006852834951132536
/usr/local/lib/python3.7/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=0.9281045751633987, f1=0.9261682242990654,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.47285604220815e-05
step: 10, loss: 2.4277074771816842e-05
step: 20, loss: 0.00016832102846819907
step: 30, loss: 0.0040268306620419025
step: 40, loss: 0.02073713205754757
step: 50, loss: 0.04359232261776924
step: 60, loss: 0.0017220836598426104
step: 70, loss: 0.0007778041181154549
step: 80, loss: 0.00045170148951001465
step: 90, loss: 0.00019493546278681606
step: 100, loss: 0.0007777256541885436
step: 110, loss: 0.16160956025123596
step: 120, loss: 0.0001044476666720584
step: 130, loss: 0.0020869714207947254
step: 140, loss: 2.6728379452833906e-05
step: 150, loss: 0.00010022400238085538
step: 160, loss: 3.932177787646651e-05
step: 170, loss: 6.998389289947227e-05
step: 180, loss: 8.223217446357012e-05
step: 190, loss: 0.002098917728289962
step: 200, loss: 0.001897266716696322
step: 210, loss: 0.0037682652473449707
step: 220, loss: 0.016426095739006996
step: 230, loss: 2.3871232770034112e-05
step: 240, loss: 0.0002702656202018261
step: 250, loss: 3.4121203498216346e-05
step: 260, loss: 3.9956958062248304e-05
step: 270, loss: 6.698981451336294e-05
```

```
step: 280, loss: 0.0022260015830397606
step: 290, loss: 5.4757379984948784e-05
step: 300, loss: 0.0008666006615385413
step: 310, loss: 8.695496944710612e-05
step: 320, loss: 0.0006952914991416037
step: 330, loss: 0.0013000387698411942
step: 340, loss: 0.0001819527824409306
step: 350, loss: 0.00029136662487871945
step: 360, loss: 0.00458285678178072
step: 370, loss: 0.0030386382713913918
step: 380, loss: 0.00033005551085807383
step: 390, loss: 0.0001195063887280412
step: 400, loss: 5.172909004613757e-05
step: 410, loss: 0.00026168531621806324
step: 420, loss: 0.00268408190459013
step: 430, loss: 0.002710238564759493
step: 440, loss: 0.0075690713711082935
step: 450, loss: 0.012135923840105534
step: 460, loss: 5.9401310863904655e-05
step: 470, loss: 0.00397678604349494
step: 480, loss: 0.005758229177445173
step: 490, loss: 0.004070581868290901
step: 500, loss: 0.0001267738116439432
step: 510, loss: 0.031442493200302124
step: 520, loss: 0.0019667872693389654
step: 530, loss: 0.0003689884324558079
/usr/local/lib/python3.7/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=0.9271402550091076, f1=0.9301688726608854,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034938151948153973
step: 10, loss: 0.00039037875831127167
step: 20, loss: 0.00010305019532097504
step: 30, loss: 0.00020559628319460899
step: 40, loss: 9.914092515828088e-05
step: 50, loss: 0.005302437115460634
step: 60, loss: 0.0012108908267691731
step: 70, loss: 0.0001370461395708844
step: 80, loss: 0.0001220771810039878
step: 90, loss: 0.0010958834318444133
step: 100, loss: 0.00023075116041582078
step: 110, loss: 0.00010521559306653216
step: 120, loss: 5.375930049922317e-05
step: 130, loss: 6.938321894267574e-05
```

```

step: 140, loss: 0.0001120713131967932
step: 150, loss: 0.007804774213582277
step: 160, loss: 0.00012899580178782344
step: 170, loss: 5.4833544709254056e-05
step: 180, loss: 7.120521331671625e-05
step: 190, loss: 0.0008709412650205195
step: 200, loss: 3.961270340369083e-05
step: 210, loss: 7.04613994457759e-05
step: 220, loss: 0.0005768905393779278
step: 230, loss: 0.0007891646237112582
step: 240, loss: 0.0013103067176416516
step: 250, loss: 4.97673281643074e-05
step: 260, loss: 8.416693890467286e-05
step: 270, loss: 6.238617061171681e-05
step: 280, loss: 3.3306518162135035e-05
step: 290, loss: 0.0003567052772268653
step: 300, loss: 6.883188325446099e-05
step: 310, loss: 0.00239356467500329
step: 320, loss: 0.0038067344576120377
step: 330, loss: 0.01077081449329853
step: 340, loss: 0.0017320533515885472
step: 350, loss: 0.00017609339556656778
step: 360, loss: 4.0535200241720304e-05
step: 370, loss: 0.00020464735280256718
step: 380, loss: 4.7901114157866687e-05
step: 390, loss: 0.005025070160627365
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 9.882983431452885e-05
step: 410, loss: 0.0018033041851595044
step: 420, loss: 0.00010830177052412182
step: 430, loss: 0.001321855466812849
step: 440, loss: 0.0001082171147572808
step: 450, loss: 0.0072687650099396706
step: 460, loss: 0.0009258179925382137
step: 470, loss: 0.0003410080971661955
step: 480, loss: 0.00014044281851965934
step: 490, loss: 0.00046376363025046885
step: 500, loss: 0.0003461756859906018
step: 510, loss: 0.0036756745539605618
step: 520, loss: 8.132825314532965e-05
step: 530, loss: 0.00043964950600638986
/usr/local/lib/python3.7/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=0.9267161410018553, f1=0.927170868347339,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002242016140371561
step: 10, loss: 8.695901487953961e-05
step: 20, loss: 9.11323368200101e-05
step: 30, loss: 3.0371244065463543e-05
step: 40, loss: 0.05223092436790466
step: 50, loss: 0.00033573273685760796
step: 60, loss: 1.441301174054388e-05
step: 70, loss: 0.0004293100500945002
step: 80, loss: 2.2976706532062963e-05
step: 90, loss: 0.00022307624749373645
step: 100, loss: 0.002826445270329714
step: 110, loss: 7.939420902403072e-05
step: 120, loss: 0.00019259807595517486
step: 130, loss: 0.027485746890306473
step: 140, loss: 0.0025013666599988937
step: 150, loss: 8.515443914802745e-05
step: 160, loss: 0.00034673407208174467
step: 170, loss: 0.00010080429638037458
step: 180, loss: 0.0002567768970038742
step: 190, loss: 0.00011259245366090909
step: 200, loss: 0.010882091708481312
step: 210, loss: 0.000398989679524675
step: 220, loss: 4.25642792833969e-05
step: 230, loss: 0.04124769940972328
step: 240, loss: 4.036070458823815e-05
step: 250, loss: 0.00010252804349875078
step: 260, loss: 0.00010162643593503162
step: 270, loss: 0.0003209835267625749
step: 280, loss: 3.3145690395031124e-05
step: 290, loss: 0.015183967538177967
step: 300, loss: 0.0010271512437611818
step: 310, loss: 0.0024001356214284897
step: 320, loss: 0.0011484628776088357
step: 330, loss: 0.008630861528217793
step: 340, loss: 0.004377080127596855
step: 350, loss: 6.797469541197643e-05
step: 360, loss: 0.00034699655952863395
step: 370, loss: 8.689043170306832e-05
step: 380, loss: 0.00011916739458683878
step: 390, loss: 0.0205985140055418
step: 400, loss: 0.00023620908905286342
step: 410, loss: 0.00012756898649968207
step: 420, loss: 0.00014942065172363073
step: 430, loss: 0.0012195236049592495
step: 440, loss: 0.00022366868506651372
step: 450, loss: 0.0017209886573255062
step: 460, loss: 0.0011480211978778243
```

```

step: 470, loss: 0.00026015829644165933
step: 480, loss: 0.0001194382130051963
step: 490, loss: 0.00043326427112333477
step: 500, loss: 0.0003127314557787031
step: 510, loss: 0.0032498128712177277
step: 520, loss: 0.0008019686792977154
step: 530, loss: 0.00019149821309838444
/usr/local/lib/python3.7/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=0.9259944495837188, f1=0.9274418604651162,
best_f1=0.9188935771214253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.4630884076468647e-05
step: 10, loss: 0.00027038113330490887
step: 20, loss: 0.00013773476530332118
step: 30, loss: 7.687291508773342e-05
step: 40, loss: 0.030466414988040924
step: 50, loss: 0.0007482119835913181
step: 60, loss: 3.979094617534429e-05
step: 70, loss: 0.001827922067604959
step: 80, loss: 3.2505478884559125e-05
step: 90, loss: 2.9841507057426497e-05
step: 100, loss: 0.0009578727767802775
step: 110, loss: 0.001284414203837514
step: 120, loss: 0.00016061053611338139
step: 130, loss: 0.12842591106891632
step: 140, loss: 0.00021117873257026076
step: 150, loss: 6.708850560244173e-05
step: 160, loss: 0.00013407852384261787
step: 170, loss: 0.0001899785565910861
step: 180, loss: 3.913562250090763e-05
step: 190, loss: 0.00031484788632951677
step: 200, loss: 6.346289592329413e-05
step: 210, loss: 0.00022487793467007577
step: 220, loss: 0.012992138043045998
step: 230, loss: 0.0001695486280368641
step: 240, loss: 0.027312031015753746
step: 250, loss: 0.00015658348274882883
step: 260, loss: 0.00012266894918866456
step: 270, loss: 0.00030605250503867865
step: 280, loss: 0.0002163829340133816
step: 290, loss: 2.1054809622000903e-05
step: 300, loss: 6.059291263227351e-05
step: 310, loss: 0.00029934506164863706
step: 320, loss: 5.298432006384246e-05

```

```

step: 330, loss: 3.9302547520492226e-05
step: 340, loss: 5.308590334607288e-05
step: 350, loss: 0.00041241623694077134
step: 360, loss: 0.0018596233567222953
step: 370, loss: 9.433556988369673e-05
step: 380, loss: 0.0003901034651789814
step: 390, loss: 0.0004909767303615808
step: 400, loss: 0.0004979443619959056
step: 410, loss: 0.00029066327260807157
step: 420, loss: 0.00038117708754725754
step: 430, loss: 0.0003527650551404804
step: 440, loss: 0.00016071289428509772
step: 450, loss: 1.7389402273693122e-05
step: 460, loss: 2.193400177930016e-05
step: 470, loss: 0.00463105458766222
step: 480, loss: 4.077377889188938e-05
step: 490, loss: 0.0003451499796938151
step: 500, loss: 0.00024244959058705717
step: 510, loss: 0.0004830869729630649
step: 520, loss: 5.553268056246452e-05
step: 530, loss: 0.00012311570753809065
/usr/local/lib/python3.7/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=0.9259436215957956, f1=0.9261553120533587,
best_f1=0.9188935771214253
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:32, 177.95it/s]
load_f1 = 0.9262564584311883
real_f1 = 0.9261176470588235
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 180.68it/s]

```