

AADatasets_bert

August 13, 2022

0.1 Install packages

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 8.48 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.5 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 19.5 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 52.4 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.4 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 33.3 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 21.7 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 | 4.37 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 26.0 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 69.4 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 7.1 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.5 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 49.8 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 60.0 MB/s
Requirement already satisfied: six>=1.5.0 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.15.0)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
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 srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 70.3 MB/s
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: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Collecting blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
    | 3.7 MB 61.4 MB/s
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 49.2 MB/s
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.7/dist-packages (from sklearn==0.0->-r requirements.txt
(line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
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 tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    |                                     | 3.0 MB 48.2 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    |                                     | 880 kB 58.5 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: 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: 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: 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: 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: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449921
sha256=a16b03e204a79f7c324d16f91e26a98405a184e1ab24bc9884efdd25e776625e
  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=77e3b014cf32699cf9f7ec9cb5bc633efc3f7a0267d150c42d9399a06dda251b
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 1.12.1+cu113
  Uninstalling torch-1.12.1+cu113:
    Successfully uninstalled torch-1.12.1+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 0.13.1+cu113 requires torch==1.12.1, but you have torch 1.4.0 which
is incompatible.

torchtext 0.13.1 requires torch==1.12.1, but you have torch 1.4.0 which is
incompatible.

torchaudio 0.12.1+cu113 requires torch==1.12.1, 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 2.7.9 requires torch<1.14,>=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 | 10.74 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-d_q8t_pq
Created temporary directory: /tmp/pip-req-tracker-stjp47m6
Initialized build tracking at /tmp/pip-req-tracker-stjp47m6
Created build tracker: /tmp/pip-req-tracker-stjp47m6
Entered build tracker: /tmp/pip-req-tracker-stjp47m6
```

```

Created temporary directory: /tmp/pip-install-bd1pyicg
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-96r5c9pa
  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-
stjp47m6'
  Running setup.py (path:/tmp/pip-req-build-96r5c9pa/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-c_19nga_
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-c_19nga_/apex.egg-info
  writing /tmp/pip-pip-egg-info-c_19nga_/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-c_19nga_/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-c_19nga_/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-c_19nga_/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-c_19nga_/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-96r5c9pa 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-stjp47m6'
Created temporary directory: /tmp/pip-unpack-3mhdsc6h
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-zc9w84z2
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-zc9w84z2
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-96r5c9pa/setup.py''';

```



```
__file__='''/tmp/pip-req-build-96r5c9pa/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-zc9w84z2
```

```
torch.__version__ = 1.4.0
```

```
running bdist_wheel
running build
running build_py
creating build
creating build/lib
creating build/lib/apex
copying apex/_autocast_utils.py -> build/lib/apex
copying apex/__init__.py -> build/lib/apex
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/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/optimizers
copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
creating build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
creating build/lib/apex/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/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
```

```

copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/_amp_state.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/utils.py -> build/lib/apex/amp
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
creating build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
creating build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
creating build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
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/xentropy
copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
creating build/lib/apex/contrib/layer_norm
copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
copying apex/contrib/layer_norm/__init__.py ->

```

```

build/lib/apex/contrib/layer_norm
    creating build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/__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
    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/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/optimizers
    copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    creating build/lib/apex/contrib/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/bottleneck
    copying apex/contrib/bottleneck/bottleneck.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/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/__init__.py ->
build/lib/apex/contrib/conv_bias_relu
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn.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/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/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_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/__init__.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
    copying
apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py ->
build/lib/apex/contrib/multihead_attn
    creating build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/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
creating 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
copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
creating build/lib/apex/transformer/layers
copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
creating build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/layers.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/mappings.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/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
creating build/lib/apex/transformer/functional
copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
creating build/lib/apex/transformer/_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/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing
copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing

```

```

    copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_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/fwd_bwd_pipelining_with_i
nterleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/__init__.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/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    installing to build/bdist.linux-x86_64/wheel
    running install
    running install_lib
    creating build/bdist.linux-x86_64
    creating build/bdist.linux-x86_64/wheel
    creating build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/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/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/optimizers
    copying build/lib/apex/optimizers/fused_adagrad.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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    creating build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    creating build/bdist.linux-x86_64/wheel/apex/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
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/contrib
    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/xentropy
    copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  copying build/lib/apex/contrib/peer_memory/__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
  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/transducer
  copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  creating build/bdist.linux-x86_64/wheel/apex/contrib/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/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
  copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
  copying build/lib/apex/contrib/optimizers/fused_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_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
  creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
  copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
  creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
  copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
  copying build/lib/apex/contrib/sparsity/permutation_search_kernels/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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck.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/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
    creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    creating build/bdist.linux-x86_64/wheel/apex/contrib/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/encdec_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn

```

```

    copying
build/lib/apex/contrib/multihead_attn/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_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating 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/_process_optimizer.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.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/wrap.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/compat.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_state.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/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/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/_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/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp

```

```

    creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/__init__.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
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    creating build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/layers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/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/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    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/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/__init__.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/distributed_test_base.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    creating build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_without_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_no_pipelining.py
-> build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  copying build/lib/apex/transformer/microbatches.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/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/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  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-zc9w84z2/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=6ff6834107e49f643eb9d85ee9d796d237d1266dadb5767b73f386266272eb12
Stored in directory: /tmp/pip-ephem-wheel-cache-
d_q8t_pq/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-stjp47m6'
/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 36.6 MB/s
Collecting awscli
  Downloading awscli-1.25.51-py3-none-any.whl (3.9 MB)
    |                                     | 3.9 MB 47.7 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                                     | 548 kB 65.2 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                                     | 79 kB 9.3 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting botocore==1.27.51
  Downloading botocore-1.27.51-py3-none-any.whl (9.0 MB)
    |                                     | 9.0 MB 58.8 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.51->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==1.27.51->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 2.7.9 requires torch<1.14,>=1.7, but you have torch 1.4.0 which is
incompatible.
Successfully installed awscli-1.25.51 botocore-1.27.51 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 16 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 0.13.1+cu113 requires torch==1.12.1, but you have torch 1.7.1 which
is incompatible.

torchtext 0.13.1 requires torch==1.12.1, but you have torch 1.7.1 which is
incompatible.

torchaudio 0.12.1+cu113 requires torch==1.12.1, 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: 1057, done.
remote: Counting objects: 100% (127/127), done.
remote: Compressing objects: 100% (95/95), done.
remote: Total 1057 (delta 52), reused 45 (delta 20), pack-reused 930
Receiving objects: 100% (1057/1057), 257.86 MiB | 17.34 MiB/s, done.
Resolving deltas: 100% (636/636), 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/AADatasets/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, 355kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 3.99MB/s]

Downloading: 100% 440M/440M [00:12<00:00, 36.0MB/s]

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

Defaults for this optimization level are:

enabled : True

opt_level : 02

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5397289991378784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5454545454545454, f1=0.5405405405405405,
best_f1=0.5405405405405405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.501953661441803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.588235294117647, f1=0.5000000000000001,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.427357941865921
/usr/local/lib/python3.7/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=0.5641025641025641, f1=0.45, best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3173300623893738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7741935483870968, f1=0.6470588235294117,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23228619992733002
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.7333333333333334,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17479249835014343
/usr/local/lib/python3.7/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=0.7741935483870968, f1=0.8387096774193549,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020641472190618515

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8484848484848484, f1=0.8750000000000001,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017082717269659042
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8750000000000001,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016796727431938052
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8750000000000001,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006164269987493753
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015272155869752169
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.002119977492839098
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001779800746589899
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008185029961168766
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8750000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002619184786453843
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8750000000000001
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, 117404.39it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7777777777777778
real_f1 = 0.7777777777777778
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, 255.45it/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 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: 0.6275483965873718

```

step: 10, loss: 0.6230438351631165
step: 20, loss: 0.22103938460350037
step: 30, loss: 0.133161723613739
step: 40, loss: 0.18410003185272217
step: 50, loss: 0.03831837326288223
step: 60, loss: 0.03510670363903046
step: 70, loss: 0.016260895878076553
step: 80, loss: 0.004897525068372488
step: 90, loss: 0.18144647777080536
step: 100, loss: 0.002086816355586052
step: 110, loss: 0.1321641504764557
step: 120, loss: 0.005281893536448479
step: 130, loss: 0.009129554964601994
step: 140, loss: 0.0015545007772743702
step: 150, loss: 0.008854984305799007
step: 160, loss: 0.0021178536117076874
step: 170, loss: 0.1148194819688797
step: 180, loss: 0.010374338366091251
step: 190, loss: 0.022530511021614075
step: 200, loss: 0.010982727631926537
step: 210, loss: 0.002872387645766139
step: 220, loss: 0.007785600610077381
step: 230, loss: 0.04298010468482971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876543209876544, f1=0.9818594104308391,
best_f1=0.9818594104308391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011843868996948004
step: 10, loss: 0.000922856735996902
step: 20, loss: 0.016146346926689148
step: 30, loss: 0.1890714168548584
step: 40, loss: 0.0700569748878479
step: 50, loss: 0.004209617618471384
step: 60, loss: 0.002672541420906782
step: 70, loss: 0.12644712626934052
step: 80, loss: 0.0027836901135742664
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.09872398525476456
step: 100, loss: 0.0038003071676939726
step: 110, loss: 0.014380901120603085

```

```

step: 120, loss: 0.001793239382095635
step: 130, loss: 0.007057523354887962
step: 140, loss: 0.0008860365487635136
step: 150, loss: 0.0012598993489518762
step: 160, loss: 0.08177220821380615
step: 170, loss: 0.0009486585040576756
step: 180, loss: 0.007805502042174339
step: 190, loss: 0.0738929882645607
step: 200, loss: 0.0012730831513181329
step: 210, loss: 0.010258035734295845
step: 220, loss: 0.08229655772447586
step: 230, loss: 0.0030689823906868696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9943502824858756, f1=0.9852440408626559,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001146308844909072
step: 10, loss: 0.0009731926838867366
step: 20, loss: 0.0005823731189593673
step: 30, loss: 0.019285595044493675
step: 40, loss: 0.15736393630504608
step: 50, loss: 0.012172703631222248
step: 60, loss: 0.002151378197595477
step: 70, loss: 0.002510198624804616
step: 80, loss: 0.0007693368825130165
step: 90, loss: 0.0029637045226991177
step: 100, loss: 0.0008551242062821984
step: 110, loss: 0.00048408537986688316
step: 120, loss: 0.018004609271883965
step: 130, loss: 0.0030087376944720745
step: 140, loss: 0.001864577759988606
step: 150, loss: 0.0013838483719155192
step: 160, loss: 0.11009683459997177
step: 170, loss: 0.0074182674288749695
step: 180, loss: 0.0036047915928065777
step: 190, loss: 0.0017663832986727357
step: 200, loss: 0.0028118137270212173
step: 210, loss: 0.0013141712406650186
step: 220, loss: 0.0006237651687115431
step: 230, loss: 0.00029233412351459265

```

```

/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9876543209876544,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030356046045199037
step: 10, loss: 0.00028803228633478284
step: 20, loss: 0.0004338346188887954
step: 30, loss: 0.00039637027657590806
step: 40, loss: 0.004357103258371353
step: 50, loss: 0.00030529650393873453
step: 60, loss: 0.000516666506882757
step: 70, loss: 0.0004935082979500294
step: 80, loss: 0.03288005292415619
step: 90, loss: 0.0005243562627583742
step: 100, loss: 0.00023672965471632779
step: 110, loss: 0.0005890087340958416
step: 120, loss: 0.002558761276304722
step: 130, loss: 0.0002578174462541938
step: 140, loss: 0.0002088890178129077
step: 150, loss: 0.21615639328956604
step: 160, loss: 0.11295878887176514
step: 170, loss: 0.003655609441921115
step: 180, loss: 0.0008882890688255429
step: 190, loss: 0.06294401735067368
step: 200, loss: 0.0023369535338133574
step: 210, loss: 0.16471855342388153
step: 220, loss: 0.0002724775404203683
step: 230, loss: 0.0014041266404092312
/usr/local/lib/python3.7/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=0.9898762654668166, f1=0.9843400447427293,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007520248182117939
step: 10, loss: 0.0008276455337181687
step: 20, loss: 0.0005053389468230307
step: 30, loss: 0.0014540092088282108
step: 40, loss: 0.0004693477530963719
step: 50, loss: 0.0004135508497711271
step: 60, loss: 0.00026066118152812123
step: 70, loss: 0.0002521080896258354

```

```

step: 80, loss: 0.0003872582456097007
step: 90, loss: 0.0001500427897553891
step: 100, loss: 0.00019612822507042438
step: 110, loss: 0.0003581676573958248
step: 120, loss: 0.00031656643841415644
step: 130, loss: 0.0022223968990147114
step: 140, loss: 0.0008480950491502881
step: 150, loss: 0.004810971673578024
step: 160, loss: 0.00042854269850067794
step: 170, loss: 0.004494400694966316
step: 180, loss: 0.0009186333045363426
step: 190, loss: 0.000562025117687881
step: 200, loss: 0.004892572294920683
step: 210, loss: 0.0001767897920217365
step: 220, loss: 0.0015430925413966179
step: 230, loss: 0.00014351749268826097
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9865168539325843,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021103539620526135
step: 10, loss: 0.0004988278378732502
step: 20, loss: 0.001123022404499352
step: 30, loss: 0.013235842809081078
step: 40, loss: 0.00015844515291973948
step: 50, loss: 0.0003686977725010365
step: 60, loss: 0.001956725027412176
step: 70, loss: 0.00019734290253836662
step: 80, loss: 0.00032089464366436005
step: 90, loss: 0.00019899608741980046
step: 100, loss: 0.02093082293868065
step: 110, loss: 0.0009133343119174242
step: 120, loss: 0.00022062835341785103
step: 130, loss: 0.0003286924911662936
step: 140, loss: 8.16326355561614e-05
step: 150, loss: 0.00039579372969456017
step: 160, loss: 0.0001992691250052303
step: 170, loss: 0.00031832282547838986
step: 180, loss: 0.023308442905545235
step: 190, loss: 0.0019352339440956712
step: 200, loss: 0.0003702173999045044
step: 210, loss: 0.006468586623668671
step: 220, loss: 0.00015174000873230398
step: 230, loss: 0.033724620938301086

```

```

/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9875706214689265,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018558402371127158
step: 10, loss: 6.826329627074301e-05
step: 20, loss: 0.00017619889695197344
step: 30, loss: 0.0004083034582436085
step: 40, loss: 9.082971519092098e-05
step: 50, loss: 0.0001282278390135616
step: 60, loss: 0.0001374075945932418
step: 70, loss: 0.013801085762679577
step: 80, loss: 0.0001062962255673483
step: 90, loss: 7.696717511862516e-05
step: 100, loss: 0.00010985812696162611
step: 110, loss: 0.0004728282510768622
step: 120, loss: 7.24322089808993e-05
step: 130, loss: 7.249112968565896e-05
step: 140, loss: 0.00010974804899888113
step: 150, loss: 0.013712973333895206
step: 160, loss: 0.04106074199080467
step: 170, loss: 0.0007232475909404457
step: 180, loss: 0.000758479000069201
step: 190, loss: 0.000139742303872481
step: 200, loss: 0.038263846188783646
step: 210, loss: 5.694157152902335e-05
step: 220, loss: 0.01662740856409073
step: 230, loss: 0.00013550405856221914
/usr/local/lib/python3.7/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=0.9899216125419933, f1=0.9866071428571428,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.255041368305683e-05
step: 10, loss: 0.00020347085956018418
step: 20, loss: 7.258875120896846e-05
step: 30, loss: 0.00015396202798001468
step: 40, loss: 0.00025698976241983473
step: 50, loss: 0.00015195510059129447
step: 60, loss: 6.053267134120688e-05
step: 70, loss: 8.793755841907114e-05

```



```

step: 80, loss: 6.349376781145111e-05
step: 90, loss: 0.001107987598516047
step: 100, loss: 7.937364716781303e-05
step: 110, loss: 0.11160381883382797
step: 120, loss: 0.08312492817640305
step: 130, loss: 8.985471504274756e-05
step: 140, loss: 7.919468043837696e-05
step: 150, loss: 0.0003060336457565427
step: 160, loss: 0.00015324993000831455
step: 170, loss: 9.105676872422919e-05
step: 180, loss: 0.00029857896151952446
step: 190, loss: 0.00016187026631087065
step: 200, loss: 7.557313074357808e-05
step: 210, loss: 9.294777555624023e-05
step: 220, loss: 0.00013183669943828136
step: 230, loss: 8.553810039302334e-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=0.9898989898989898, f1=0.9844097995545658,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.556282798759639e-05
step: 10, loss: 0.0001279987336602062
step: 20, loss: 0.00027661569765768945
step: 30, loss: 0.00010763008322101086
step: 40, loss: 0.025034479796886444
step: 50, loss: 4.496150359045714e-05
step: 60, loss: 0.0005104415467940271
step: 70, loss: 9.40369427553378e-05
step: 80, loss: 5.742784924223088e-05
step: 90, loss: 8.332230208907276e-05
step: 100, loss: 0.00012881359725724906
step: 110, loss: 5.427750511444174e-05
step: 120, loss: 3.652900704764761e-05
step: 130, loss: 4.305865877540782e-05
step: 140, loss: 4.022762368549593e-05
step: 150, loss: 0.0003352052590344101
step: 160, loss: 0.000151893706060946
step: 170, loss: 6.772277265554294e-05
step: 180, loss: 0.0001365308853564784
step: 190, loss: 0.00015177424938883632
step: 200, loss: 5.47745221410878e-05
step: 210, loss: 7.173584162956104e-05
step: 220, loss: 7.884304068284109e-05
step: 230, loss: 4.9592534196563065e-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=0.9910313901345291, f1=0.9844444444444443,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004050726711284369
step: 10, loss: 3.531071342877112e-05
step: 20, loss: 5.157968553248793e-05
step: 30, loss: 6.637213664362207e-05
step: 40, loss: 5.8643352531362325e-05
step: 50, loss: 3.508332883939147e-05
step: 60, loss: 0.00018180193728767335
step: 70, loss: 5.8314315538154915e-05
step: 80, loss: 4.214976797811687e-05
step: 90, loss: 0.0001007301762001589
step: 100, loss: 4.61709605588112e-05
step: 110, loss: 4.704611637862399e-05
step: 120, loss: 7.56403460400179e-05
step: 130, loss: 4.0964896470541134e-05
step: 140, loss: 0.032107025384902954
step: 150, loss: 0.013423687778413296
step: 160, loss: 2.5990428184741177e-05
step: 170, loss: 5.9726244217017666e-05
step: 180, loss: 4.0145518141798675e-05
step: 190, loss: 0.00837087631225586
step: 200, loss: 6.148874672362581e-05
step: 210, loss: 3.940569149563089e-05
step: 220, loss: 5.2205228712409735e-05
step: 230, loss: 7.216211088234559e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9910313901345291, f1=0.9844097995545658,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.18879535573069e-05
step: 10, loss: 5.166421760804951e-05
step: 20, loss: 9.337087976746261e-05
step: 30, loss: 0.014984820038080215
step: 40, loss: 0.0002434427588013932
step: 50, loss: 4.369007729110308e-05
step: 60, loss: 7.953184103826061e-05
step: 70, loss: 9.632293949835002e-05

```

```

step: 80, loss: 0.00011663988698273897
step: 90, loss: 0.010309970937669277
step: 100, loss: 5.245824286248535e-05
step: 110, loss: 0.0600246861577034
step: 120, loss: 2.3491160391131416e-05
step: 130, loss: 3.368917532498017e-05
step: 140, loss: 2.7614718419499695e-05
step: 150, loss: 0.03416028246283531
step: 160, loss: 3.513502451824024e-05
step: 170, loss: 0.020307334139943123
step: 180, loss: 5.738788240705617e-05
step: 190, loss: 3.143630237900652e-05
step: 200, loss: 0.00027640911866910756
step: 210, loss: 3.556152296368964e-05
step: 220, loss: 3.439412103034556e-05
step: 230, loss: 0.00011251000250922516
/usr/local/lib/python3.7/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=0.9910313901345291, f1=0.9865470852017937,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9026537958998233e-05
step: 10, loss: 3.977430969825946e-05
step: 20, loss: 2.6783898647408932e-05
step: 30, loss: 5.6261975259985775e-05
step: 40, loss: 5.69279000046663e-05
step: 50, loss: 0.00020093715284019709
step: 60, loss: 8.938344399211928e-05
step: 70, loss: 4.907112815999426e-05
step: 80, loss: 0.0001161159307230264
step: 90, loss: 2.6385358069092035e-05
step: 100, loss: 2.3274975319509394e-05
step: 110, loss: 4.746601916849613e-05
step: 120, loss: 6.320480315480381e-05
step: 130, loss: 2.6981564587913454e-05
step: 140, loss: 0.016929106786847115
step: 150, loss: 4.765954508911818e-05
step: 160, loss: 3.4654887713259086e-05
step: 170, loss: 3.381264832569286e-05
step: 180, loss: 0.030149739235639572
step: 190, loss: 5.8933248510584235e-05
step: 200, loss: 1.8447206457494758e-05
step: 210, loss: 2.559211316111032e-05
step: 220, loss: 2.219104499090463e-05
step: 230, loss: 0.04136693477630615

```

```

/usr/local/lib/python3.7/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=0.9910313901345291, f1=0.9854423292273236,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.714252908248454e-05
step: 10, loss: 4.3162039219168946e-05
step: 20, loss: 8.839698421070352e-05
step: 30, loss: 7.031349377939478e-05
step: 40, loss: 2.7335312552168034e-05
step: 50, loss: 5.494834840646945e-05
step: 60, loss: 2.8002164981444366e-05
step: 70, loss: 1.8581396943773143e-05
step: 80, loss: 2.5837869543465786e-05
step: 90, loss: 3.685560295707546e-05
step: 100, loss: 1.6145155314006843e-05
step: 110, loss: 0.023634860292077065
step: 120, loss: 0.024315625429153442
step: 130, loss: 0.0001404039649059996
step: 140, loss: 2.6445279218023643e-05
step: 150, loss: 4.66073070128914e-05
step: 160, loss: 3.773854405153543e-05
step: 170, loss: 4.039176201331429e-05
step: 180, loss: 2.8095355446566828e-05
step: 190, loss: 4.462153810891323e-05
step: 200, loss: 4.300480577512644e-05
step: 210, loss: 1.9471757696010172e-05
step: 220, loss: 2.2287946194410324e-05
step: 230, loss: 2.589750147308223e-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=0.9910313901345291, f1=0.987709497206704,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.571882214397192e-05
step: 10, loss: 0.01586528867483139
step: 20, loss: 2.9711733077419922e-05
step: 30, loss: 3.272549292887561e-05
step: 40, loss: 0.011227162554860115
step: 50, loss: 2.9637683837790973e-05
step: 60, loss: 2.8392938475008123e-05
step: 70, loss: 3.247928907512687e-05

```

```

step: 80, loss: 2.5640660169301555e-05
step: 90, loss: 3.022590135515202e-05
step: 100, loss: 2.3912130927783437e-05
step: 110, loss: 4.9290305469185114e-05
step: 120, loss: 1.7102580386563204e-05
step: 130, loss: 2.4225120796472766e-05
step: 140, loss: 4.053638258483261e-05
step: 150, loss: 2.9391398129519075e-05
step: 160, loss: 0.0006873080274090171
step: 170, loss: 2.3841106667532586e-05
step: 180, loss: 3.5153287171851844e-05
step: 190, loss: 7.51721890992485e-05
step: 200, loss: 2.1103314793435857e-05
step: 210, loss: 3.529210516717285e-05
step: 220, loss: 4.153494592173956e-05
step: 230, loss: 9.300086094299331e-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=0.9910313901345291, f1=0.987709497206704,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.9441953554633074e-05
step: 10, loss: 1.513193728897022e-05
step: 20, loss: 2.8266755180084147e-05
step: 30, loss: 3.91270405089017e-05
step: 40, loss: 0.0001526091800769791
step: 50, loss: 0.046249739825725555
step: 60, loss: 0.00016797777789179236
step: 70, loss: 3.073942934861407e-05
step: 80, loss: 4.413359420141205e-05
step: 90, loss: 0.00016634601342957467
step: 100, loss: 2.350236536585726e-05
step: 110, loss: 3.24885368172545e-05
step: 120, loss: 3.086659489781596e-05
step: 130, loss: 0.025294199585914612
step: 140, loss: 2.054467586276587e-05
step: 150, loss: 3.5202359867980704e-05
step: 160, loss: 2.4153741833288223e-05
step: 170, loss: 2.619870429043658e-05
step: 180, loss: 8.409217116422951e-05
step: 190, loss: 0.00014736753655597568
step: 200, loss: 2.4768929506535642e-05
step: 210, loss: 3.121641930192709e-05
step: 220, loss: 7.432923302985728e-05
step: 230, loss: 2.1356665456551127e-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=0.9910313901345291, f1=0.987709497206704,
best_f1=0.9852440408626559
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, 196.16it/s]
load_f1 = 0.9943502824858756
real_f1 = 0.992108229988726
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 242.43it/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, 414kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 4.07MB/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: 0.62013179063797
step: 10, loss: 0.5266956090927124
step: 20, loss: 0.37027060985565186
step: 30, loss: 0.0561799630522728
step: 40, loss: 0.09413570910692215
step: 50, loss: 0.040527451783418655
step: 60, loss: 0.021504998207092285
step: 70, loss: 0.06510265171527863
step: 80, loss: 0.050710346549749374
step: 90, loss: 0.12217589467763901
step: 100, loss: 0.12840817868709564
step: 110, loss: 0.04071688652038574
step: 120, loss: 0.0827401876449585
step: 130, loss: 0.10504821687936783
step: 140, loss: 0.07466708868741989
step: 150, loss: 0.022939911112189293
step: 160, loss: 0.014659238047897816
step: 170, loss: 0.13020136952400208
step: 180, loss: 0.06600774079561234
step: 190, loss: 0.016633575782179832
step: 200, loss: 0.15628433227539062
step: 210, loss: 0.09521941840648651
step: 220, loss: 0.14024998247623444

```



```

step: 230, loss: 0.1581711620092392
step: 240, loss: 0.07470930367708206
step: 250, loss: 0.013157385401427746
step: 260, loss: 0.10840889811515808
step: 270, loss: 0.007724026218056679
step: 280, loss: 0.03701817989349365
step: 290, loss: 0.06479309499263763
step: 300, loss: 0.033916402608156204
step: 310, loss: 0.1346523016691208
step: 320, loss: 0.13298270106315613
step: 330, loss: 0.009106838144361973
step: 340, loss: 0.03181936964392662
step: 350, loss: 0.012704429216682911
step: 360, loss: 0.022998174652457237
step: 370, loss: 0.09483228623867035
step: 380, loss: 0.007831281051039696
step: 390, loss: 0.11966981738805771
step: 400, loss: 0.21759140491485596
step: 410, loss: 0.04025975242257118
step: 420, loss: 0.02402578294277191
step: 430, loss: 0.1491602063179016
step: 440, loss: 0.013715121895074844
step: 450, loss: 0.0018372912891209126
step: 460, loss: 0.003592773573473096
step: 470, loss: 0.12925294041633606
step: 480, loss: 0.03539527580142021
step: 490, loss: 0.15248934924602509
step: 500, loss: 0.03814462199807167
step: 510, loss: 0.03495926782488823
step: 520, loss: 0.05813293904066086
step: 530, loss: 0.004116517957299948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9443430656934306, f1=0.9455040871934605,
best_f1=0.9455040871934605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29515039920806885
step: 10, loss: 0.05491148680448532
step: 20, loss: 0.012023831717669964
step: 30, loss: 0.00694246357306838
step: 40, loss: 0.07701259851455688

```

step: 50, loss: 0.09465977549552917
step: 60, loss: 0.004286894109100103
step: 70, loss: 0.03107849322259426
step: 80, loss: 0.010496048256754875
step: 90, loss: 0.019432533532381058
step: 100, loss: 0.005503418855369091
step: 110, loss: 0.011965729296207428
step: 120, loss: 0.03595053777098656
step: 130, loss: 0.11900138109922409
step: 140, loss: 0.09552767872810364
step: 150, loss: 0.04048550873994827
step: 160, loss: 0.0035777969751507044
step: 170, loss: 0.028672685846686363
step: 180, loss: 0.1012209877371788
step: 190, loss: 0.06718708574771881
step: 200, loss: 0.0018638648325577378
step: 210, loss: 0.02803543210029602
step: 220, loss: 0.04355313628911972
step: 230, loss: 0.007151440251618624
step: 240, loss: 0.0041659241542220116
step: 250, loss: 0.007734611164778471
step: 260, loss: 0.0019460771000012755
step: 270, loss: 0.06970027834177017
step: 280, loss: 0.0031469673849642277
step: 290, loss: 0.007300565950572491
step: 300, loss: 0.13557998836040497
step: 310, loss: 0.011337156407535076
step: 320, loss: 0.07261698693037033
step: 330, loss: 0.030134519562125206
step: 340, loss: 0.007248419336974621
step: 350, loss: 0.0063117872923612595
step: 360, loss: 0.07669049501419067
step: 370, loss: 0.05239682272076607
step: 380, loss: 0.02861977554857731
step: 390, loss: 0.03080470860004425
step: 400, loss: 0.01775750145316124
step: 410, loss: 0.014163283631205559
step: 420, loss: 0.051439426839351654
step: 430, loss: 0.017156941816210747
step: 440, loss: 0.15506979823112488
step: 450, loss: 0.007179485633969307
step: 460, loss: 0.003943036776036024
step: 470, loss: 0.055463071912527084
step: 480, loss: 0.24154818058013916
step: 490, loss: 0.008435426279902458
step: 500, loss: 0.2000914067029953
step: 510, loss: 0.0054017165675759315
step: 520, loss: 0.0639198049902916

```

step: 530, loss: 0.002289279131218791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9482999534233814, f1=0.9479553903345725,
best_f1=0.9479553903345725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04966103285551071
step: 10, loss: 0.07412800937891006
step: 20, loss: 0.04583192244172096
step: 30, loss: 0.20900464057922363
step: 40, loss: 0.0053270356729626656
step: 50, loss: 0.0555124431848526
step: 60, loss: 0.013119773007929325
step: 70, loss: 0.003676737193018198
step: 80, loss: 0.0013354438124224544
step: 90, loss: 0.008828964084386826
step: 100, loss: 0.056911908090114594
step: 110, loss: 0.015497470274567604
step: 120, loss: 0.010900674387812614
step: 130, loss: 0.0021943196188658476
step: 140, loss: 0.0093425577506423
step: 150, loss: 0.06079995632171631
step: 160, loss: 0.02074933797121048
step: 170, loss: 0.05874191224575043
step: 180, loss: 0.012225082144141197
step: 190, loss: 0.004627819173038006
step: 200, loss: 0.00561896339058876
step: 210, loss: 0.0726623460650444
step: 220, loss: 0.041526321321725845
step: 230, loss: 0.09464067965745926
step: 240, loss: 0.0009257701458409429
step: 250, loss: 0.04649434983730316
step: 260, loss: 0.01043193880468607
step: 270, loss: 0.002544532995671034
step: 280, loss: 0.14921236038208008
step: 290, loss: 0.0044837836176157
step: 300, loss: 0.03812059015035629
step: 310, loss: 0.1597190797328949
step: 320, loss: 0.006038372404873371
step: 330, loss: 0.0024857947137206793
step: 340, loss: 0.0011689838720485568

```

```

step: 350, loss: 0.003995475824922323
step: 360, loss: 0.011583710089325905
step: 370, loss: 0.008790036663413048
step: 380, loss: 0.008413342759013176
step: 390, loss: 0.0027262310031801462
step: 400, loss: 0.010493526235222816
step: 410, loss: 0.003980245441198349
step: 420, loss: 0.07717260718345642
step: 430, loss: 0.004983188584446907
step: 440, loss: 0.0043066805228590965
step: 450, loss: 0.08443595468997955
step: 460, loss: 0.0770607739686966
step: 470, loss: 0.08316581696271896
step: 480, loss: 0.00790680106729269
step: 490, loss: 0.0033046358730643988
step: 500, loss: 0.005809354595839977
step: 510, loss: 0.001889285515062511
step: 520, loss: 0.2997545599937439
step: 530, loss: 0.025498777627944946
/usr/local/lib/python3.7/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=0.9428838951310862, f1=0.9425502101821578,
best_f1=0.9479553903345725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025785140693187714
step: 10, loss: 0.013315172865986824
step: 20, loss: 0.001751226489432156
step: 30, loss: 0.002616593847051263
step: 40, loss: 0.0010048998519778252
step: 50, loss: 0.0008646288770250976
step: 60, loss: 0.0007721782894805074
step: 70, loss: 0.0008394901524297893
step: 80, loss: 0.012372750788927078
step: 90, loss: 0.030381502583622932
step: 100, loss: 0.002994716865941882
step: 110, loss: 0.012729895301163197
step: 120, loss: 0.000262717338046059
step: 130, loss: 0.051458101719617844
step: 140, loss: 0.006317502819001675
step: 150, loss: 0.0017665241612121463
step: 160, loss: 0.026502063497900963
step: 170, loss: 0.0029123711865395308
step: 180, loss: 0.06936497241258621
step: 190, loss: 0.11175260692834854
step: 200, loss: 0.004708949942141771

```

```

step: 210, loss: 0.042878396809101105
step: 220, loss: 0.023578057065606117
step: 230, loss: 0.09403256326913834
step: 240, loss: 0.006315185688436031
step: 250, loss: 0.02159028872847557
step: 260, loss: 0.0017356290481984615
step: 270, loss: 0.005973262712359428
step: 280, loss: 0.07029405981302261
step: 290, loss: 0.018078841269016266
step: 300, loss: 0.0029302285984158516
step: 310, loss: 0.037894245237112045
step: 320, loss: 0.04096253588795662
step: 330, loss: 0.029328858479857445
step: 340, loss: 0.012099655345082283
step: 350, loss: 0.021538937464356422
step: 360, loss: 0.023484904319047928
step: 370, loss: 0.003827976994216442
step: 380, loss: 0.0011370540596544743
step: 390, loss: 0.0007888803957030177
step: 400, loss: 0.0013212530175223947
step: 410, loss: 0.0006355738150887191
step: 420, loss: 0.0026242113672196865
step: 430, loss: 0.18248920142650604
step: 440, loss: 0.004274290055036545
step: 450, loss: 0.0021019645500928164
step: 460, loss: 0.010172762908041477
step: 470, loss: 0.02210129424929619
step: 480, loss: 0.1091768816113472
step: 490, loss: 0.012634813785552979
step: 500, loss: 0.0016712355427443981
step: 510, loss: 0.020507827401161194
step: 520, loss: 0.14891700446605682
step: 530, loss: 0.02469581738114357
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9488982653539615, f1=0.9437148217636022,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011580348014831543
step: 10, loss: 0.04824145883321762
step: 20, loss: 0.10487474501132965

```

step: 30, loss: 0.0020349333062767982
step: 40, loss: 0.003484955755993724
step: 50, loss: 0.0018962655449286103
step: 60, loss: 0.028925105929374695
step: 70, loss: 0.0004169030871707946
step: 80, loss: 0.0004115184710826725
step: 90, loss: 0.00940252747386694
step: 100, loss: 0.03422774001955986
step: 110, loss: 0.0005640678573399782
step: 120, loss: 0.000605124922003597
step: 130, loss: 0.001537490519694984
step: 140, loss: 0.00042280531488358974
step: 150, loss: 0.022731300443410873
step: 160, loss: 0.002878311090171337
step: 170, loss: 0.04848426952958107
step: 180, loss: 0.00029988394817337394
step: 190, loss: 0.0014835594920441508
step: 200, loss: 0.07885190844535828
step: 210, loss: 0.005472562741488218
step: 220, loss: 0.0033426634036004543
step: 230, loss: 0.008572385646402836
step: 240, loss: 0.003167891874909401
step: 250, loss: 0.0002923008578363806
step: 260, loss: 0.007541862316429615
step: 270, loss: 0.00017094302165787667
step: 280, loss: 0.00026132629136554897
step: 290, loss: 0.13162215054035187
step: 300, loss: 0.0013279335107654333
step: 310, loss: 0.010634198784828186
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 320, loss: 0.11760580539703369
step: 330, loss: 0.008936424739658833
step: 340, loss: 0.004342122934758663
step: 350, loss: 0.0026396699249744415
step: 360, loss: 0.005604122299700975
step: 370, loss: 0.009519786573946476
step: 380, loss: 0.00035288528306409717
step: 390, loss: 0.00019713792426045984
step: 400, loss: 0.0043101380579173565
step: 410, loss: 0.00021587277296930552
step: 420, loss: 0.0017558584222570062
step: 430, loss: 0.0013770083896815777
step: 440, loss: 0.0643695518374443
step: 450, loss: 0.03598390147089958
step: 460, loss: 0.0007517413469031453
step: 470, loss: 0.005337602458894253
step: 480, loss: 0.0013620562385767698
step: 490, loss: 0.0003750407777260989

```
step: 500, loss: 0.01056871097534895
step: 510, loss: 0.019885774701833725
step: 520, loss: 0.005415683146566153
step: 530, loss: 0.010168248787522316
/usr/local/lib/python3.7/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=0.942615239887112, f1=0.9405152224824357,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001315549947321415
step: 10, loss: 0.0006698189536109567
step: 20, loss: 0.07195451855659485
step: 30, loss: 0.0021892196964472532
step: 40, loss: 0.05499947816133499
step: 50, loss: 0.021985094994306564
step: 60, loss: 0.0005379768554121256
step: 70, loss: 0.0005241857143118978
step: 80, loss: 0.011897461488842964
step: 90, loss: 0.00047015585005283356
step: 100, loss: 0.0019750534556806087
step: 110, loss: 0.0015662038931623101
step: 120, loss: 0.0003748997114598751
step: 130, loss: 0.0005643467884510756
step: 140, loss: 0.018123172223567963
step: 150, loss: 0.0011654166737571359
step: 160, loss: 0.0008110946509987116
step: 170, loss: 0.0008185566985048354
step: 180, loss: 0.0015613068826496601
step: 190, loss: 0.000711719156242907
step: 200, loss: 0.0002215634158346802
step: 210, loss: 0.0030977940186858177
step: 220, loss: 0.0017313692951574922
step: 230, loss: 0.002114372095093131
step: 240, loss: 0.01638944074511528
step: 250, loss: 0.053103893995285034
step: 260, loss: 0.000367190659744665
step: 270, loss: 0.11717428267002106
step: 280, loss: 0.002453719498589635
step: 290, loss: 0.0010878031607717276
step: 300, loss: 0.0024400914553552866
step: 310, loss: 0.0035608590114861727
step: 320, loss: 0.0005992280202917755
step: 330, loss: 0.00021172197011765093
step: 340, loss: 0.12738248705863953
step: 350, loss: 0.021209681406617165
```

```

step: 360, loss: 0.007594974245876074
step: 370, loss: 0.0033383984118700027
step: 380, loss: 0.00012792447523679584
step: 390, loss: 0.012097734026610851
step: 400, loss: 0.0001080860965885222
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 410, loss: 0.000623390544205904
step: 420, loss: 0.00843411311507225
step: 430, loss: 0.0005237903678789735
step: 440, loss: 7.87461904110387e-05
step: 450, loss: 0.000492305145598948
step: 460, loss: 0.00023602432338520885
step: 470, loss: 0.015747610479593277
step: 480, loss: 0.0006571430712938309
step: 490, loss: 0.014845896512269974
step: 500, loss: 0.001793479430489242
step: 510, loss: 0.020595012232661247
step: 520, loss: 0.005902876146137714
step: 530, loss: 0.0042636278085410595
/usr/local/lib/python3.7/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=0.9468283582089553, f1=0.942271880819367,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016767489723861217
step: 10, loss: 0.0006857140106149018
step: 20, loss: 0.00015468316269107163
step: 30, loss: 0.00025328484480269253
step: 40, loss: 0.00038693612441420555
step: 50, loss: 0.008205204270780087
step: 60, loss: 0.00036251291749067605
step: 70, loss: 0.00023888531723059714
step: 80, loss: 0.009534628130495548
step: 90, loss: 0.0001015753805404529
step: 100, loss: 0.0006044225883670151
step: 110, loss: 0.009328865446150303
step: 120, loss: 0.0042557548731565475
step: 130, loss: 0.000390605884604156
step: 140, loss: 0.00015158178575802594
step: 150, loss: 0.00027710021822713315
step: 160, loss: 0.00011251337855355814
step: 170, loss: 9.547785884933546e-05
step: 180, loss: 0.0015647668624296784
step: 190, loss: 0.0012544598430395126
step: 200, loss: 0.0021514769177883863

```



```

step: 210, loss: 0.00020740910258609802
step: 220, loss: 0.00015320313104894012
step: 230, loss: 0.04009731486439705
step: 240, loss: 0.0002486305311322212
step: 250, loss: 0.0008480356773361564
step: 260, loss: 0.0005086185992695391
step: 270, loss: 0.017698563635349274
step: 280, loss: 0.016965528950095177
step: 290, loss: 0.00022746081231161952
step: 300, loss: 0.0016361699672415853
step: 310, loss: 0.00012222191435284913
step: 320, loss: 7.8741111792624e-05
step: 330, loss: 0.0010759781580418348
step: 340, loss: 0.04948704317212105
step: 350, loss: 0.0004018380423076451
step: 360, loss: 0.0017791531281545758
step: 370, loss: 0.0003803610452450812
step: 380, loss: 0.0008388595888391137
step: 390, loss: 0.003929476719349623
step: 400, loss: 0.003102395683526993
step: 410, loss: 0.002933796960860491
step: 420, loss: 0.0013545661931857467
step: 430, loss: 8.457233343506232e-05
step: 440, loss: 0.026647282764315605
step: 450, loss: 0.0025455926079303026
step: 460, loss: 0.000888710783328861
step: 470, loss: 0.0006224556127563119
step: 480, loss: 0.006904931738972664
step: 490, loss: 0.00016352906823158264
step: 500, loss: 8.486305159749463e-05
step: 510, loss: 0.0004390841058921069
step: 520, loss: 0.006529036909341812
step: 530, loss: 0.014268084429204464
/usr/local/lib/python3.7/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=0.9486575600565238, f1=0.9471715755025713,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0062646628357470036
step: 10, loss: 0.0075058769434690475
step: 20, loss: 4.9929953092942014e-05
step: 30, loss: 7.341925811488181e-05
step: 40, loss: 0.0001352944236714393
step: 50, loss: 0.0068758283741772175
step: 60, loss: 0.004374154843389988

```

step: 70, loss: 6.12255753367208e-05
step: 80, loss: 0.08759155869483948
step: 90, loss: 0.0004910668940283358
step: 100, loss: 6.966336513869464e-05
step: 110, loss: 0.00023434331524185836
step: 120, loss: 0.14978134632110596
step: 130, loss: 0.0001956860942300409
step: 140, loss: 0.010323286056518555
step: 150, loss: 0.0005448854644782841
step: 160, loss: 0.02497202157974243
step: 170, loss: 0.1387440711259842
step: 180, loss: 0.00024757604114711285
step: 190, loss: 0.0033349154982715845
step: 200, loss: 0.0014032552717253566
step: 210, loss: 0.0018676223699003458
step: 220, loss: 0.0021579123567789793
step: 230, loss: 0.015699604526162148
step: 240, loss: 0.0002561259607318789
step: 250, loss: 8.736624295124784e-05
step: 260, loss: 9.896158735500649e-05
step: 270, loss: 0.002553927944973111
step: 280, loss: 0.03141956403851509
step: 290, loss: 0.0003813495277427137
step: 300, loss: 0.0009601420606486499
step: 310, loss: 0.00629683630540967
step: 320, loss: 0.0007654414512217045
step: 330, loss: 0.0032951468601822853
step: 340, loss: 0.00012075569975422695
step: 350, loss: 7.683194417040795e-05
step: 360, loss: 0.0017370145069435239
step: 370, loss: 0.000748803315218538
step: 380, loss: 0.0058382293209433556
step: 390, loss: 0.13488635420799255
step: 400, loss: 0.00019901491759810597
step: 410, loss: 0.0006892182864248753
step: 420, loss: 0.002050739247351885
step: 430, loss: 0.06447742879390717
step: 440, loss: 0.029314914718270302
step: 450, loss: 0.0008951341733336449
step: 460, loss: 0.003542139660567045
step: 470, loss: 0.0006270802696235478
step: 480, loss: 0.0004656141682062298
step: 490, loss: 0.007544007617980242
step: 500, loss: 0.01797170750796795
step: 510, loss: 0.0005293563008308411
step: 520, loss: 0.0025045450311154127
step: 530, loss: 0.05679420754313469
/usr/local/lib/python3.7/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=0.946927374301676, f1=0.9463459759481962,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063576712273061275
step: 10, loss: 0.0004320750304032117
step: 20, loss: 0.0015763654373586178
step: 30, loss: 0.0012504023034125566
step: 40, loss: 0.0003779480466619134
step: 50, loss: 0.00020334604778327048
step: 60, loss: 0.0001677142718108371
step: 70, loss: 0.0004795839777216315
step: 80, loss: 0.0009262370294891298
step: 90, loss: 0.00021677483164239675
step: 100, loss: 0.0004163614066783339
step: 110, loss: 0.00012624265218619257
step: 120, loss: 0.00019458784663584083
step: 130, loss: 0.006501108407974243
step: 140, loss: 0.00016593262262176722
step: 150, loss: 0.0001340062590315938
step: 160, loss: 0.0016460097394883633
step: 170, loss: 0.0495966412127018
step: 180, loss: 0.0001528074499219656
step: 190, loss: 0.00024277193006128073
step: 200, loss: 0.00010232090426143259
step: 210, loss: 0.0010940415086224675
step: 220, loss: 0.00034876709105446935
step: 230, loss: 0.07232896983623505
step: 240, loss: 0.00018624593212734908
step: 250, loss: 0.00013896566815674305
step: 260, loss: 0.004918599035590887
step: 270, loss: 5.788297494291328e-05
step: 280, loss: 0.00719714118167758
step: 290, loss: 0.04933852329850197
step: 300, loss: 0.00015274879115168005
step: 310, loss: 0.018615597859025
step: 320, loss: 0.001284956349991262
step: 330, loss: 0.00020605554163921624
step: 340, loss: 0.0001852386922109872
step: 350, loss: 0.000367369590094313
step: 360, loss: 4.1028753912542015e-05
step: 370, loss: 0.0007272193324752152
step: 380, loss: 6.479317380581051e-05
step: 390, loss: 5.7960474805440754e-05
step: 400, loss: 7.785701745888218e-05

```

```

step: 410, loss: 0.005643269047141075
step: 420, loss: 6.688952271360904e-05
step: 430, loss: 8.687309309607372e-05
step: 440, loss: 0.001804774859920144
step: 450, loss: 8.725519001018256e-05
step: 460, loss: 6.260853842832148e-05
step: 470, loss: 5.762365253758617e-05
step: 480, loss: 0.00031325436430051923
step: 490, loss: 0.0005915026413276792
step: 500, loss: 0.002561215776950121
step: 510, loss: 0.000747250858694315
step: 520, loss: 0.001490596099756658
step: 530, loss: 0.0003287880390416831
/usr/local/lib/python3.7/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=0.9441340782122905, f1=0.9480519480519481,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022034107241779566
step: 10, loss: 0.00013621716061607003
step: 20, loss: 5.118070839671418e-05
step: 30, loss: 3.796990495175123e-05
step: 40, loss: 0.0002101723657688126
step: 50, loss: 0.0024373745545744896
step: 60, loss: 0.0008375574834644794
step: 70, loss: 3.358588583068922e-05
step: 80, loss: 7.542337698396295e-05
step: 90, loss: 0.00015173164138104767
step: 100, loss: 5.086285818833858e-05
step: 110, loss: 3.054925036849454e-05
step: 120, loss: 5.9960937505820766e-05
step: 130, loss: 7.222597196232527e-05
step: 140, loss: 0.025552863255143166
step: 150, loss: 0.009108426980674267
step: 160, loss: 5.254794450593181e-05
step: 170, loss: 0.00011546481982804835
step: 180, loss: 6.756971561117098e-05
step: 190, loss: 4.111427551833913e-05
step: 200, loss: 0.008472265675663948
step: 210, loss: 8.091816562227905e-05
step: 220, loss: 0.0770692229270935
step: 230, loss: 0.0007069972925819457
step: 240, loss: 7.498950435547158e-05
step: 250, loss: 4.032402648590505e-05
step: 260, loss: 0.001385265844874084

```

```

step: 270, loss: 0.0006114780553616583
step: 280, loss: 0.0029825642704963684
step: 290, loss: 3.482264946796931e-05
step: 300, loss: 4.8016398068284616e-05
step: 310, loss: 4.351222014520317e-05
step: 320, loss: 0.00134106760378927
step: 330, loss: 0.0003515855350997299
step: 340, loss: 0.004096126649528742
step: 350, loss: 2.840838351403363e-05
step: 360, loss: 8.847384742693976e-05
step: 370, loss: 0.002436497015878558
step: 380, loss: 0.007071792613714933
step: 390, loss: 0.00024954768014140427
step: 400, loss: 0.0017119792755693197
step: 410, loss: 0.002117063384503126
step: 420, loss: 6.74567636451684e-05
step: 430, loss: 6.715217023156583e-05
step: 440, loss: 0.0009865803876891732
step: 450, loss: 0.0022384326439350843
step: 460, loss: 4.409712346387096e-05
step: 470, loss: 8.492997585562989e-05
step: 480, loss: 5.690216494258493e-05
step: 490, loss: 7.618342351634055e-05
step: 500, loss: 0.0014858996728435159
step: 510, loss: 0.00022147373238112777
step: 520, loss: 5.260267062112689e-05
step: 530, loss: 0.000861102482303977
/usr/local/lib/python3.7/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=0.9479117785077428, f1=0.9519813519813519,
best_f1=0.9437148217636022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000320659251883626
step: 10, loss: 8.286988304462284e-05
step: 20, loss: 3.4558055631350726e-05
step: 30, loss: 0.0001752061361912638
step: 40, loss: 5.9389425587141886e-05
step: 50, loss: 9.670052531873807e-05
step: 60, loss: 0.02810821495950222
step: 70, loss: 0.021813485771417618
step: 80, loss: 3.211781586287543e-05
step: 90, loss: 3.722501060110517e-05
step: 100, loss: 0.0031528242398053408
step: 110, loss: 6.60354780848138e-05
step: 120, loss: 2.1408752218121663e-05

```

```

step: 130, loss: 5.0068389100488275e-05
step: 140, loss: 0.014848719350993633
step: 150, loss: 2.7260159185971133e-05
step: 160, loss: 0.00011524560977704823
step: 170, loss: 4.372839612187818e-05
step: 180, loss: 6.558858149219304e-05
step: 190, loss: 0.00042434060014784336
step: 200, loss: 0.004522304516285658
step: 210, loss: 0.00012993530253879726
step: 220, loss: 0.0018496176926419139
step: 230, loss: 4.862627247348428e-05
step: 240, loss: 0.0006550055113621056
step: 250, loss: 3.6919071135343984e-05
step: 260, loss: 3.544102582964115e-05
step: 270, loss: 8.447004074696451e-05
step: 280, loss: 0.00013480751658789814
step: 290, loss: 0.00012315994536038488
step: 300, loss: 0.0012324448907747865
step: 310, loss: 9.931487147696316e-05
step: 320, loss: 7.525279943365604e-05
step: 330, loss: 0.00045253103598952293
step: 340, loss: 3.951642065658234e-05
step: 350, loss: 0.0003071195096708834
step: 360, loss: 3.504884443827905e-05
step: 370, loss: 0.0007790034869685769
step: 380, loss: 0.0003026259655598551
step: 390, loss: 3.6854573409073055e-05
step: 400, loss: 0.0005923677235841751
step: 410, loss: 0.0003114143037237227
step: 420, loss: 0.009272223338484764
step: 430, loss: 3.608134647947736e-05
step: 440, loss: 0.00020874195615760982
step: 450, loss: 0.00047041321522556245
step: 460, loss: 0.00022538438497576863
step: 470, loss: 0.00016712263459339738
step: 480, loss: 0.0015132336411625147
step: 490, loss: 0.0037161093205213547
step: 500, loss: 7.724213355686516e-05
step: 510, loss: 4.162261029705405e-05
step: 520, loss: 3.786168235819787e-05
step: 530, loss: 3.937655492336489e-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=0.949810606060606, f1=0.9467232437529468,  
best_f1=0.9467232437529468  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 3.415897663217038e-05  
step: 10, loss: 2.1732972527388483e-05  
step: 20, loss: 5.0570237362990156e-05  
step: 30, loss: 0.001601369003765285  
step: 40, loss: 0.0004400309408083558  
step: 50, loss: 0.037556808441877365  
step: 60, loss: 0.0015830590855330229  
step: 70, loss: 3.391359496163204e-05  
step: 80, loss: 0.00016113521996885538  
step: 90, loss: 0.0002640527964103967  
step: 100, loss: 6.564764771610498e-05  
step: 110, loss: 4.1650375351309776e-05  
step: 120, loss: 3.0384850106202066e-05  
step: 130, loss: 0.0007295522373169661  
step: 140, loss: 2.5152472517220303e-05  
step: 150, loss: 2.436273280181922e-05  
step: 160, loss: 2.096184834954329e-05  
step: 170, loss: 0.00010043555812444538  
step: 180, loss: 0.0016858342569321394  
step: 190, loss: 5.3720443247584626e-05  
step: 200, loss: 4.5928707550046965e-05  
step: 210, loss: 6.210213905433193e-05  
step: 220, loss: 4.8523750592721626e-05  
step: 230, loss: 2.4090655642794445e-05  
step: 240, loss: 6.821969873271883e-05  
step: 250, loss: 3.1376828701468185e-05  
step: 260, loss: 7.818251469871029e-05  
step: 270, loss: 9.212614531861618e-05  
step: 280, loss: 0.0012701600790023804  
step: 290, loss: 0.0003174970916006714  
step: 300, loss: 0.0008263427880592644  
step: 310, loss: 0.0002458324597682804  
step: 320, loss: 3.4859585866797715e-05  
step: 330, loss: 5.281611083773896e-05  
step: 340, loss: 3.0125549528747797e-05  
step: 350, loss: 0.0022137807682156563  
step: 360, loss: 0.029137546196579933  
step: 370, loss: 0.0036310136783868074  
step: 380, loss: 0.00016119454812724143  
step: 390, loss: 0.0003232947492506355  
step: 400, loss: 4.083654494024813e-05  
step: 410, loss: 0.00015812183846719563  
step: 420, loss: 0.0032358733005821705
```

```

step: 430, loss: 0.024268222972750664
step: 440, loss: 0.023306166753172874
step: 450, loss: 5.434226477518678e-05
step: 460, loss: 6.480658339569345e-05
step: 470, loss: 2.192284955526702e-05
step: 480, loss: 0.0014697767328470945
step: 490, loss: 0.0003009800857398659
step: 500, loss: 1.8912995074060746e-05
step: 510, loss: 0.000469283782877028
step: 520, loss: 2.636324461491313e-05
step: 530, loss: 3.523916893755086e-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=0.9484440315838365, f1=0.9490740740740741,
best_f1=0.9467232437529468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002574762562289834
step: 10, loss: 2.7424624931882136e-05
step: 20, loss: 2.2514883312396705e-05
step: 30, loss: 0.00042631776886992157
step: 40, loss: 3.68547516700346e-05
step: 50, loss: 3.173062577843666e-05
step: 60, loss: 3.16553887387272e-05
step: 70, loss: 6.567202945007011e-05
step: 80, loss: 7.706634642090648e-05
step: 90, loss: 0.029635004699230194
step: 100, loss: 0.00026653148233890533
step: 110, loss: 0.0002700344775803387
step: 120, loss: 0.00025564542738720775
step: 130, loss: 2.829639197443612e-05
step: 140, loss: 0.0004398248565848917
step: 150, loss: 0.00028694033971987665
step: 160, loss: 3.773013668251224e-05
step: 170, loss: 0.00015073968097567558
step: 180, loss: 0.003595322836190462
step: 190, loss: 2.5737430405570194e-05
step: 200, loss: 2.5066825401154347e-05
step: 210, loss: 4.2152212699875236e-05
step: 220, loss: 1.8577775335870683e-05
step: 230, loss: 2.7237680114922114e-05
step: 240, loss: 8.848679135553539e-05
step: 250, loss: 2.2649086531600915e-05
step: 260, loss: 2.6060666641569696e-05
step: 270, loss: 1.8451091818860732e-05
step: 280, loss: 2.1960177036817186e-05

```



```

step: 290, loss: 1.6785923435236327e-05
step: 300, loss: 2.4794902856228873e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 310, loss: 0.001529017579741776
step: 320, loss: 0.0028239069506525993
step: 330, loss: 0.0015907802153378725
step: 340, loss: 2.233605664514471e-05
step: 350, loss: 4.2082887375727296e-05
step: 360, loss: 1.7322372514172457e-05
step: 370, loss: 4.843595888814889e-05
step: 380, loss: 0.0006256456836126745
step: 390, loss: 0.00014866159472148865
step: 400, loss: 3.885668775183149e-05
step: 410, loss: 0.0013242758577689528
step: 420, loss: 1.8410131815471686e-05
step: 430, loss: 3.288824518676847e-05
step: 440, loss: 2.5536011889926158e-05
step: 450, loss: 3.8342583138728514e-05
step: 460, loss: 0.001776977675035596
step: 470, loss: 3.220703365514055e-05
step: 480, loss: 0.0012112518306821585
step: 490, loss: 1.9162640455760993e-05
step: 500, loss: 2.5155219191219658e-05
step: 510, loss: 3.43352512572892e-05
step: 520, loss: 0.006304512731730938
step: 530, loss: 1.814925781218335e-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=0.9503480278422275, f1=0.9509713228492137,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009083914919756353
step: 10, loss: 1.5951503883115947e-05
step: 20, loss: 1.7482572729932144e-05
step: 30, loss: 2.115559072990436e-05
step: 40, loss: 6.7482927988749e-05
step: 50, loss: 8.644288027426228e-05
step: 60, loss: 2.9635803002747707e-05
step: 70, loss: 2.2291544155450538e-05
step: 80, loss: 3.266125713707879e-05
step: 90, loss: 2.108841908921022e-05

```

```
step: 100, loss: 2.6646013793651946e-05
step: 110, loss: 2.8491882403614e-05
step: 120, loss: 1.689017517492175e-05
step: 130, loss: 0.021708091720938683
step: 140, loss: 0.0022816092241555452
step: 150, loss: 1.3619514902529772e-05
step: 160, loss: 0.0026693157851696014
step: 170, loss: 9.818329999689013e-05
step: 180, loss: 1.7877324353321455e-05
step: 190, loss: 7.839711906854063e-05
step: 200, loss: 7.42468037060462e-05
step: 210, loss: 2.3937986043165438e-05
step: 220, loss: 1.8618697140482254e-05
step: 230, loss: 2.2082995201344602e-05
step: 240, loss: 5.1329549023648724e-05
step: 250, loss: 2.1144342099432833e-05
step: 260, loss: 0.001489712274633348
step: 270, loss: 1.723663262964692e-05
step: 280, loss: 1.6350120858987793e-05
step: 290, loss: 0.0008340433705598116
step: 300, loss: 0.00011217241262784228
step: 310, loss: 2.429930100333877e-05
step: 320, loss: 0.0007546833367086947
step: 330, loss: 3.836030009551905e-05
step: 340, loss: 5.361361036193557e-05
step: 350, loss: 3.211712100892328e-05
step: 360, loss: 0.01194994430989027
step: 370, loss: 2.502161260053981e-05
step: 380, loss: 2.472748383297585e-05
step: 390, loss: 0.028195645660161972
step: 400, loss: 9.844360465649515e-05
step: 410, loss: 1.3109168321534526e-05
step: 420, loss: 0.00011570449714781716
step: 430, loss: 5.125857569510117e-05
step: 440, loss: 0.00016624924319330603
step: 450, loss: 3.7919227906968445e-05
step: 460, loss: 0.0003270115703344345
step: 470, loss: 1.9754666936933063e-05
step: 480, loss: 2.5118342819041573e-05
step: 490, loss: 2.170649531763047e-05
step: 500, loss: 1.3936150935478508e-05
step: 510, loss: 0.00034583982778713107
step: 520, loss: 0.0006985960062593222
step: 530, loss: 0.04642412066459656
/usr/local/lib/python3.7/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=0.949041608228144, f1=0.946927374301676,
```

```
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.115522460779175e-05
step: 10, loss: 1.948275530594401e-05
step: 20, loss: 4.6684664994245395e-05
step: 30, loss: 0.0004102828388568014
step: 40, loss: 0.08015831559896469
step: 50, loss: 4.200197145109996e-05
step: 60, loss: 1.4401796761376318e-05
step: 70, loss: 0.017603516578674316
step: 80, loss: 2.3073760530678555e-05
step: 90, loss: 1.6771060472819954e-05
step: 100, loss: 0.0004049217386636883
step: 110, loss: 0.000817279564216733
step: 120, loss: 0.0004034301673527807
step: 130, loss: 0.11512455344200134
step: 140, loss: 4.992128015146591e-05
step: 150, loss: 1.929301470227074e-05
step: 160, loss: 5.692517515853979e-05
step: 170, loss: 1.9602048269007355e-05
step: 180, loss: 1.4446532986767124e-05
step: 190, loss: 0.00018353544874116778
step: 200, loss: 1.995587081182748e-05
step: 210, loss: 0.0009947309736162424
step: 220, loss: 1.5370320397778414e-05
step: 230, loss: 0.0004978022770956159
step: 240, loss: 2.142361154255923e-05
step: 250, loss: 1.5098421499715187e-05
step: 260, loss: 3.0064420570852235e-05
step: 270, loss: 2.172522908949759e-05
step: 280, loss: 0.00014936363731976599
step: 290, loss: 1.3854220014763996e-05
step: 300, loss: 1.9445318685029633e-05
step: 310, loss: 2.683957245608326e-05
step: 320, loss: 0.00010123928223038092
step: 330, loss: 2.563213638495654e-05
step: 340, loss: 1.4699841813126113e-05
step: 350, loss: 1.5660838471376337e-05
step: 360, loss: 0.0018832519417628646
step: 370, loss: 1.643572613829747e-05
step: 380, loss: 2.144975405826699e-05
step: 390, loss: 2.9760120014543645e-05
step: 400, loss: 2.493945612513926e-05
step: 410, loss: 1.685296228970401e-05
step: 420, loss: 0.0010429295944049954
step: 430, loss: 4.0859405999071896e-05
```

```

step: 440, loss: 0.000139879688504152
step: 450, loss: 1.502016675658524e-05
step: 460, loss: 5.8357298257760704e-05
step: 470, loss: 0.002202274277806282
step: 480, loss: 1.3008594578423072e-05
step: 490, loss: 1.4684931556985248e-05
step: 500, loss: 4.324115070630796e-05
step: 510, loss: 2.011609285546001e-05
step: 520, loss: 1.6033458450692706e-05
step: 530, loss: 0.0003653730673249811
/usr/local/lib/python3.7/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=0.946927374301676, f1=0.9485396383866481,
best_f1=0.9509713228492137
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 [00:22, 250.89it/s]
load_f1 = 0.949438202247191
real_f1 = 0.9474671669793621

```

```

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, 247.80it/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: 0.5521329045295715
step: 10, loss: 0.37764233350753784
step: 20, loss: 0.3925336003303528
step: 30, loss: 0.32667088508605957
step: 40, loss: 0.16401493549346924
step: 50, loss: 0.4055115878582001
step: 60, loss: 0.1893320530653
step: 70, loss: 0.14819195866584778
step: 80, loss: 0.21720591187477112
step: 90, loss: 0.2947181463241577
step: 100, loss: 0.34851858019828796
step: 110, loss: 0.1568726748228073
step: 120, loss: 0.12298138439655304
step: 130, loss: 0.16072148084640503
step: 140, loss: 0.21044927835464478
step: 150, loss: 0.17993566393852234
step: 160, loss: 0.2757418155670166
step: 170, loss: 0.22539086639881134
step: 180, loss: 0.0645548552274704
step: 190, loss: 0.2531631290912628

```

```

step: 200, loss: 0.2569867968559265
step: 210, loss: 0.21734273433685303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6751361161524501, f1=0.6953271028037382,
best_f1=0.6953271028037382
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10221831500530243
step: 10, loss: 0.19590508937835693
step: 20, loss: 0.23240689933300018
step: 30, loss: 0.21139481663703918
step: 40, loss: 0.3145654797554016
step: 50, loss: 0.19560857117176056
step: 60, loss: 0.3261101543903351
step: 70, loss: 0.11497249454259872
step: 80, loss: 0.12112371623516083
step: 90, loss: 0.13909779489040375
step: 100, loss: 0.0020733713172376156
step: 110, loss: 0.12381435930728912
step: 120, loss: 0.13637176156044006
step: 130, loss: 0.022139225155115128
step: 140, loss: 0.15384434163570404
step: 150, loss: 0.21765224635601044
step: 160, loss: 0.10949785262346268
step: 170, loss: 0.1319807916879654
step: 180, loss: 0.12480128556489944
step: 190, loss: 0.1300259530544281
step: 200, loss: 0.07104004919528961
step: 210, loss: 0.11904889345169067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7099567099567099, f1=0.6826086956521739,
best_f1=0.6826086956521739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.054031111299999161
step: 10, loss: 0.15947702527046204
step: 20, loss: 0.0571853443980217
step: 30, loss: 0.11815212666988373
step: 40, loss: 0.1341419219970703
step: 50, loss: 0.05881698429584503
step: 60, loss: 0.14191068708896637
step: 70, loss: 0.10529045760631561
step: 80, loss: 0.1050821840763092
step: 90, loss: 0.0327739492058754
step: 100, loss: 0.23996268212795258
step: 110, loss: 0.14706525206565857
step: 120, loss: 0.1663854718208313
step: 130, loss: 0.13867801427841187
step: 140, loss: 0.2376229465007782
step: 150, loss: 0.19332444667816162
step: 160, loss: 0.013882217928767204
step: 170, loss: 0.09953170269727707
step: 180, loss: 0.1159643903374672
step: 190, loss: 0.2787829041481018
step: 200, loss: 0.03359793499112129
step: 210, loss: 0.08293597400188446
/usr/local/lib/python3.7/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=0.7080979284369116, f1=0.7342256214149139,
best_f1=0.6826086956521739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11581069976091385
step: 10, loss: 0.03940252959728241
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.0954497829079628
step: 30, loss: 0.16880638897418976
step: 40, loss: 0.01868700236082077
step: 50, loss: 0.23963214457035065
step: 60, loss: 0.11551974713802338
step: 70, loss: 0.14518634974956512
step: 80, loss: 0.029876427724957466
step: 90, loss: 0.021005915477871895
step: 100, loss: 0.2939850091934204
step: 110, loss: 0.13282334804534912
step: 120, loss: 0.1157456487417221
step: 130, loss: 0.28195658326148987
step: 140, loss: 0.13417983055114746
step: 150, loss: 0.045583099126815796
step: 160, loss: 0.050697147846221924

```



```

step: 170, loss: 0.15685899555683136
step: 180, loss: 0.24052956700325012
step: 190, loss: 0.09368972480297089
step: 200, loss: 0.29262053966522217
step: 210, loss: 0.13640055060386658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7407407407407407, f1=0.7333333333333334,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12098389118909836
step: 10, loss: 0.21938695013523102
step: 20, loss: 0.4140418767929077
step: 30, loss: 0.12707456946372986
step: 40, loss: 0.16709405183792114
step: 50, loss: 0.1947459876537323
step: 60, loss: 0.15028215944766998
step: 70, loss: 0.15278466045856476
step: 80, loss: 0.036536067724227905
step: 90, loss: 0.15206503868103027
step: 100, loss: 0.005745209753513336
step: 110, loss: 0.26002490520477295
step: 120, loss: 0.15899188816547394
step: 130, loss: 0.03611224889755249
step: 140, loss: 0.040205102413892746
step: 150, loss: 0.14892855286598206
step: 160, loss: 0.1109752207994461
step: 170, loss: 0.05849379673600197
step: 180, loss: 0.09166701138019562
step: 190, loss: 0.01108507439494133
step: 200, loss: 0.08589563518762589
step: 210, loss: 0.027403900399804115
/usr/local/lib/python3.7/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=0.7289719626168224, f1=0.7265774378585086,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.042392704635858536

```

```

step: 10, loss: 0.08913897722959518
step: 20, loss: 0.06671051681041718
step: 30, loss: 0.0017262311885133386
step: 40, loss: 0.029982388019561768
step: 50, loss: 0.027620023116469383
step: 60, loss: 0.08932743221521378
step: 70, loss: 0.05076588690280914
step: 80, loss: 0.147136390209198
step: 90, loss: 0.07925713062286377
step: 100, loss: 0.0054817222990095615
step: 110, loss: 0.013647756539285183
step: 120, loss: 0.1114615947008133
step: 130, loss: 0.09704720973968506
step: 140, loss: 0.08306793868541718
step: 150, loss: 0.03217359259724617
step: 160, loss: 0.013382477685809135
step: 170, loss: 0.10685358941555023
step: 180, loss: 0.03180166333913803
step: 190, loss: 0.13703566789627075
step: 200, loss: 0.011407462880015373
step: 210, loss: 0.03663209080696106
/usr/local/lib/python3.7/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=0.7265625, f1=0.7246963562753036, best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021339861676096916
step: 10, loss: 0.12641891837120056
step: 20, loss: 0.010976821184158325
step: 30, loss: 0.004169470630586147
step: 40, loss: 0.02797897346317768
step: 50, loss: 0.08131254464387894
step: 60, loss: 0.015578437596559525
step: 70, loss: 0.01450183242559433
step: 80, loss: 0.02384406141936779
step: 90, loss: 0.06465267390012741
step: 100, loss: 0.011738221161067486
step: 110, loss: 0.22934755682945251
step: 120, loss: 0.15631137788295746
step: 130, loss: 0.09556937962770462
step: 140, loss: 0.045512303709983826
step: 150, loss: 0.24722054600715637
step: 160, loss: 0.011808091774582863
step: 170, loss: 0.04072904214262962
step: 180, loss: 0.183695450425148
step: 190, loss: 0.06570851802825928

```

```

step: 200, loss: 0.02445336803793907
step: 210, loss: 0.055918339639902115
/usr/local/lib/python3.7/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=0.732824427480916, f1=0.731610337972167,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04936617612838745
step: 10, loss: 0.1860600858926773
step: 20, loss: 0.016325879842042923
step: 30, loss: 0.055011481046676636
step: 40, loss: 0.03590654209256172
step: 50, loss: 0.14110712707042694
step: 60, loss: 0.2017340511083603
step: 70, loss: 0.008968206122517586
step: 80, loss: 0.05222701281309128
step: 90, loss: 0.014830841682851315
step: 100, loss: 0.04362039268016815
step: 110, loss: 0.04392387345433235
step: 120, loss: 0.039140306413173676
step: 130, loss: 0.046198103576898575
step: 140, loss: 0.03285490348935127
step: 150, loss: 0.01756139099597931
step: 160, loss: 0.05558512359857559
step: 170, loss: 0.004322608467191458
step: 180, loss: 0.16435496509075165
step: 190, loss: 0.05745457857847214
step: 200, loss: 0.013378364033997059
step: 210, loss: 0.21967744827270508
/usr/local/lib/python3.7/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=0.736196319018405, f1=0.711018711018711,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02065819688141346
step: 10, loss: 0.028854070231318474
step: 20, loss: 0.0008456601062789559
step: 30, loss: 0.0014598775887861848
step: 40, loss: 0.0600740872323513
step: 50, loss: 0.1975422203540802
step: 60, loss: 0.13949620723724365
step: 70, loss: 0.0063922591507434845

```

```

step: 80, loss: 0.12728191912174225
step: 90, loss: 0.002402221318334341
step: 100, loss: 0.039044979959726334
step: 110, loss: 0.014832945540547371
step: 120, loss: 0.005744958762079477
step: 130, loss: 0.11562526226043701
step: 140, loss: 0.04772181436419487
step: 150, loss: 0.08378233760595322
step: 160, loss: 0.009012525901198387
step: 170, loss: 0.025995168834924698
step: 180, loss: 0.04716577008366585
step: 190, loss: 0.002491658553481102
step: 200, loss: 0.02515866607427597
step: 210, loss: 0.0084008714184165
/usr/local/lib/python3.7/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=0.7308447937131631, f1=0.7054108216432866,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09172205626964569
step: 10, loss: 0.28090474009513855
step: 20, loss: 0.0013716338435187936
step: 30, loss: 0.08802443742752075
step: 40, loss: 0.0007528096903115511
step: 50, loss: 0.005884500220417976
step: 60, loss: 0.1370626837015152
step: 70, loss: 0.11955664306879044
step: 80, loss: 0.04820481687784195
step: 90, loss: 0.050674039870500565
step: 100, loss: 0.03219357877969742
step: 110, loss: 0.01755695976316929
step: 120, loss: 0.04009290784597397
step: 130, loss: 0.0645039826631546
step: 140, loss: 0.014111353084445
step: 150, loss: 0.06871774047613144
step: 160, loss: 0.03197067603468895
step: 170, loss: 0.0033613150008022785
step: 180, loss: 0.06351824849843979
step: 190, loss: 0.06147946044802666
step: 200, loss: 0.012799946591258049
step: 210, loss: 0.03325763717293739
/usr/local/lib/python3.7/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=0.7210626185958253, f1=0.7129094412331405,

```

```

best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08243518322706223
step: 10, loss: 0.020847227424383163
step: 20, loss: 0.10813082009553909
step: 30, loss: 0.000859511608723551
step: 40, loss: 0.027305930852890015
step: 50, loss: 0.03244984149932861
step: 60, loss: 0.0888751745223999
step: 70, loss: 0.011648250743746758
step: 80, loss: 0.06736616790294647
step: 90, loss: 0.008337805978953838
step: 100, loss: 0.03415893018245697
step: 110, loss: 0.018674399703741074
step: 120, loss: 0.08321691304445267
step: 130, loss: 0.018691997975111008
step: 140, loss: 0.000807782809715718
step: 150, loss: 0.0024949435610324144
step: 160, loss: 0.023924851790070534
step: 170, loss: 0.01445272471755743
step: 180, loss: 0.007768855895847082
step: 190, loss: 0.00842093676328659
step: 200, loss: 0.0019518511835485697
step: 210, loss: 0.003972125239670277
/usr/local/lib/python3.7/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=0.7294589178356714, f1=0.7228915662650602,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05699041858315468
step: 10, loss: 0.002181281568482518
step: 20, loss: 0.016759095713496208
step: 30, loss: 0.07722385227680206
step: 40, loss: 0.23062331974506378
step: 50, loss: 0.001630971091799438
step: 60, loss: 0.037808842957019806
step: 70, loss: 0.0038054294418543577
step: 80, loss: 0.0028620073571801186
step: 90, loss: 0.06203709915280342
step: 100, loss: 0.030970998108386993
step: 110, loss: 0.008466193452477455
step: 120, loss: 0.06929437071084976
step: 130, loss: 0.009616725146770477

```

```

step: 140, loss: 0.01872778683900833
step: 150, loss: 0.07244149595499039
step: 160, loss: 0.0210952777415514
step: 170, loss: 0.023960599675774574
step: 180, loss: 0.00840920303016901
step: 190, loss: 0.0021320756059139967
step: 200, loss: 0.0016619190573692322
step: 210, loss: 0.019876955077052116
/usr/local/lib/python3.7/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=0.7310606060606061, f1=0.7151051625239006,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05373011901974678
step: 10, loss: 0.0021370970644056797
step: 20, loss: 0.06649370491504669
step: 30, loss: 0.20202255249023438
step: 40, loss: 0.03680182993412018
step: 50, loss: 0.051903825253248215
step: 60, loss: 0.008166447281837463
step: 70, loss: 0.049029819667339325
step: 80, loss: 0.013481516391038895
step: 90, loss: 0.0023725859355181456
step: 100, loss: 0.10329803079366684
step: 110, loss: 0.0013493680162355304
step: 120, loss: 0.05201977491378784
step: 130, loss: 0.12781091034412384
step: 140, loss: 0.03393426537513733
step: 150, loss: 0.09190795570611954
step: 160, loss: 0.015303620137274265
step: 170, loss: 0.018900172784924507
step: 180, loss: 0.21088039875030518
step: 190, loss: 0.0011700111208483577
step: 200, loss: 0.0012657004408538342
step: 210, loss: 0.019402673467993736
/usr/local/lib/python3.7/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=0.725868725868726, f1=0.7090558766859344,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012305212439969182
step: 10, loss: 0.03671539947390556

```

```

step: 20, loss: 0.014204471372067928
step: 30, loss: 0.0035086951684206724
step: 40, loss: 0.003571033477783203
step: 50, loss: 0.03129016235470772
step: 60, loss: 0.004291725344955921
step: 70, loss: 0.0118480883538723
step: 80, loss: 0.052958402782678604
step: 90, loss: 0.13959313929080963
step: 100, loss: 0.02311699651181698
step: 110, loss: 0.008414039388298988
step: 120, loss: 0.047328267246484756
step: 130, loss: 0.0030009187757968903
step: 140, loss: 0.05474619194865227
step: 150, loss: 0.0046896496787667274
step: 160, loss: 0.008619776926934719
step: 170, loss: 0.03061777725815773
step: 180, loss: 0.0011885948479175568
step: 190, loss: 0.0053342655301094055
step: 200, loss: 0.0007010935223661363
step: 210, loss: 0.010598720982670784
/usr/local/lib/python3.7/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=0.7231404958677686, f1=0.7167381974248928,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011294604279100895
step: 10, loss: 0.00017638018471188843
step: 20, loss: 0.002247675321996212
step: 30, loss: 0.036470748484134674
step: 40, loss: 0.0023288349620997906
step: 50, loss: 0.006609112955629826
step: 60, loss: 0.018517274409532547
step: 70, loss: 0.0013246734160929918
step: 80, loss: 0.0011061852565035224
step: 90, loss: 0.0026808595284819603
step: 100, loss: 0.003732360666617751
step: 110, loss: 0.003976201638579369
step: 120, loss: 0.04520266875624657
step: 130, loss: 0.0029198909178376198
step: 140, loss: 0.0004806879151146859
step: 150, loss: 0.0018222964135929942
step: 160, loss: 0.0011062389239668846
step: 170, loss: 0.0011343236546963453
step: 180, loss: 0.00014348106924444437
step: 190, loss: 0.006845844443887472

```

```

step: 200, loss: 0.026961732655763626
step: 210, loss: 0.03136647120118141
/usr/local/lib/python3.7/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=0.7268907563025211, f1=0.7221006564551422,
best_f1=0.7333333333333334
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, 448.41it/s]
load_f1 = 0.7398843930635838
real_f1 = 0.7413127413127413
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 243.22it/s]

```


1.1.5 Walmart-Amazon - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5773806571960449
step: 10, loss: 0.36834716796875
step: 20, loss: 0.28300046920776367
step: 30, loss: 0.4306255280971527
step: 40, loss: 0.4462270140647888
step: 50, loss: 0.2883056700229645
step: 60, loss: 0.21026571094989777
step: 70, loss: 0.25748714804649353
step: 80, loss: 0.31195905804634094
step: 90, loss: 0.26060599088668823
step: 100, loss: 0.31054484844207764
step: 110, loss: 0.29147055745124817
step: 120, loss: 0.07026802003383636
step: 130, loss: 0.13857048749923706
step: 140, loss: 0.13415710628032684
step: 150, loss: 0.18776538968086243
step: 160, loss: 0.034935928881168365
step: 170, loss: 0.16143010556697845
step: 180, loss: 0.012233014218509197
step: 190, loss: 0.1507890522480011
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7135678391959798, f1=0.7518796992481204,
best_f1=0.7518796992481204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25162774324417114
step: 10, loss: 0.15097497403621674
step: 20, loss: 0.19355307519435883
step: 30, loss: 0.08940991759300232
step: 40, loss: 0.09929046779870987
step: 50, loss: 0.06335774809122086
step: 60, loss: 0.23266616463661194
step: 70, loss: 0.11760253459215164
step: 80, loss: 0.13844658434391022
step: 90, loss: 0.11224698275327682
step: 100, loss: 0.013407351449131966
step: 110, loss: 0.057996053248643875
step: 120, loss: 0.23700858652591705
step: 130, loss: 0.10230179876089096
step: 140, loss: 0.1606789380311966
step: 150, loss: 0.07894951105117798
step: 160, loss: 0.06077978014945984
step: 170, loss: 0.041190244257450104
step: 180, loss: 0.12705084681510925
step: 190, loss: 0.03169409930706024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8063660477453581, f1=0.8125, best_f1=0.8125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 0, loss: 0.07521460950374603
step: 10, loss: 0.2013920247554779
step: 20, loss: 0.0781974345445633
step: 30, loss: 0.011796039529144764
step: 40, loss: 0.03957528620958328
step: 50, loss: 0.07765477895736694
step: 60, loss: 0.1749875843524933
step: 70, loss: 0.1583194136619568
step: 80, loss: 0.03871257230639458
step: 90, loss: 0.11138477176427841

```

```

step: 100, loss: 0.03214993327856064
step: 110, loss: 0.013469408266246319
step: 120, loss: 0.09073085337877274
step: 130, loss: 0.006602954119443893
step: 140, loss: 0.020492658019065857
step: 150, loss: 0.1330258697271347
step: 160, loss: 0.05060270428657532
step: 170, loss: 0.08332648128271103
step: 180, loss: 0.0478217676281929
step: 190, loss: 0.10770028084516525
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8222811671087534, f1=0.8263157894736842,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0629919096827507
step: 10, loss: 0.3535459339618683
step: 20, loss: 0.019204143434762955
step: 30, loss: 0.018727265298366547
step: 40, loss: 0.08352146297693253
step: 50, loss: 0.05570422112941742
step: 60, loss: 0.016080724075436592
step: 70, loss: 0.007061978802084923
step: 80, loss: 0.07898519933223724
step: 90, loss: 0.02909093350172043
step: 100, loss: 0.03690207004547119
step: 110, loss: 0.0021383038256317377
step: 120, loss: 0.07928235828876495
step: 130, loss: 0.21500276029109955
step: 140, loss: 0.0072807203978300095
step: 150, loss: 0.0036323608364909887
step: 160, loss: 0.004095037002116442
step: 170, loss: 0.06810717284679413
step: 180, loss: 0.009816519916057587
step: 190, loss: 0.11692378669977188
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.8044077134986225, f1=0.7944444444444444,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.25001099705696106
step: 10, loss: 0.034286823123693466
step: 20, loss: 0.005469129886478186
step: 30, loss: 0.008287522941827774
step: 40, loss: 0.15750257670879364
step: 50, loss: 0.1433546543121338
step: 60, loss: 0.11671128123998642
step: 70, loss: 0.022566698491573334
step: 80, loss: 0.030705569311976433
step: 90, loss: 0.04866928234696388
step: 100, loss: 0.0032113022170960903
step: 110, loss: 0.009412629529833794
step: 120, loss: 0.10425910353660583
step: 130, loss: 0.054230011999607086
step: 140, loss: 0.050221383571624756
step: 150, loss: 0.026060549542307854
step: 160, loss: 0.12683767080307007
step: 170, loss: 0.06570345163345337
step: 180, loss: 0.012327312491834164
step: 190, loss: 0.15542584657669067
```

```
/usr/local/lib/python3.7/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=0.8010752688172043, f1=0.8130081300813009,
best_f1=0.8263157894736842
```

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

An input tensor was not cuda.

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

```
step: 0, loss: 0.01936735399067402
step: 10, loss: 0.21895171701908112
step: 20, loss: 0.002544767688959837
step: 30, loss: 0.00126373372040689
step: 40, loss: 0.17070479691028595
step: 50, loss: 0.06985792517662048
step: 60, loss: 0.08068203926086426
step: 70, loss: 0.024619465693831444
step: 80, loss: 0.022678261622786522
step: 90, loss: 0.002358793281018734
step: 100, loss: 0.0017537602689117193
step: 110, loss: 0.0058142999187111855
step: 120, loss: 0.06788506358861923
step: 130, loss: 0.02005528099834919
step: 140, loss: 0.10624825209379196
step: 150, loss: 0.007457963656634092
step: 160, loss: 0.05954239144921303
step: 170, loss: 0.05852564424276352
```

```

step: 180, loss: 0.1435410976409912
step: 190, loss: 0.027322590351104736
/usr/local/lib/python3.7/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=0.789873417721519, f1=0.8041775456919059,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0045628598891198635
step: 10, loss: 0.006627195049077272
step: 20, loss: 0.04688519611954689
step: 30, loss: 0.014567182399332523
step: 40, loss: 0.0028600560035556555
step: 50, loss: 0.03923066705465317
step: 60, loss: 0.011144815012812614
step: 70, loss: 0.0032096190843731165
step: 80, loss: 0.029782922938466072
step: 90, loss: 0.02785695157945156
step: 100, loss: 0.05457054451107979
step: 110, loss: 0.0030026868917047977
step: 120, loss: 0.004575433675199747
step: 130, loss: 0.11509738117456436
step: 140, loss: 0.0076079098507761955
step: 150, loss: 0.012037784792482853
step: 160, loss: 0.11977863311767578
step: 170, loss: 0.039817001670598984
step: 180, loss: 0.0015470818616449833
step: 190, loss: 0.018417440354824066
/usr/local/lib/python3.7/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=0.8062015503875969, f1=0.8188976377952756,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021274950355291367
step: 10, loss: 0.0038967770524322987
step: 20, loss: 0.017564047127962112
step: 30, loss: 0.06407716125249863
step: 40, loss: 0.009031274355947971
step: 50, loss: 0.010563108138740063
step: 60, loss: 0.0009078051662072539
step: 70, loss: 0.035543061792850494
step: 80, loss: 0.0006470202351920307
step: 90, loss: 0.009464956820011139

```

```

step: 100, loss: 0.006572023034095764
step: 110, loss: 0.00206855614669621
step: 120, loss: 0.0017025466077029705
step: 130, loss: 0.03482190519571304
step: 140, loss: 0.0006300419336184859
step: 150, loss: 0.0023065004497766495
step: 160, loss: 0.0021467567421495914
step: 170, loss: 0.006839352194219828
step: 180, loss: 0.009376213885843754
step: 190, loss: 0.009251796640455723
/usr/local/lib/python3.7/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=0.8219178082191781, f1=0.8166666666666668,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012272198218852282
step: 10, loss: 0.02428155392408371
step: 20, loss: 0.0016660999972373247
step: 30, loss: 0.0016181602841243148
step: 40, loss: 0.012055057100951672
step: 50, loss: 0.0005412494647316635
step: 60, loss: 0.0033752054441720247
step: 70, loss: 0.0006789767649024725
step: 80, loss: 0.0009526147623546422
step: 90, loss: 0.03853940963745117
step: 100, loss: 0.016850214451551437
step: 110, loss: 0.0006356313824653625
step: 120, loss: 0.0004895238671451807
step: 130, loss: 0.00045160058652982116
step: 140, loss: 0.04194045811891556
step: 150, loss: 0.0023777077440172434
step: 160, loss: 0.0006436606636270881
step: 170, loss: 0.12526506185531616
step: 180, loss: 0.004072041716426611
step: 190, loss: 0.002812328515574336
/usr/local/lib/python3.7/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=0.80719794344473, f1=0.8241469816272966,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009242057800292969
step: 10, loss: 0.022799955680966377

```

```

step: 20, loss: 0.02766740322113037
step: 30, loss: 0.0014389341231435537
step: 40, loss: 0.035102810710668564
step: 50, loss: 0.0019335465040057898
step: 60, loss: 0.003004200989380479
step: 70, loss: 0.0011523206485435367
step: 80, loss: 0.004323354922235012
step: 90, loss: 0.0013355750124901533
step: 100, loss: 0.0010687126778066158
step: 110, loss: 0.0013818576699122787
step: 120, loss: 0.18937760591506958
step: 130, loss: 0.000541743531357497
step: 140, loss: 0.0009453992242924869
step: 150, loss: 0.004053564742207527
step: 160, loss: 0.11369800567626953
step: 170, loss: 0.0031566452234983444
step: 180, loss: 0.0007400171016342938
step: 190, loss: 0.017270179465413094
/usr/local/lib/python3.7/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=0.8031914893617021, f1=0.8099173553719008,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001591612701304257
step: 10, loss: 0.0010071367723867297
step: 20, loss: 0.1682939976453781
step: 30, loss: 0.012810979038476944
step: 40, loss: 0.0009177444153465331
step: 50, loss: 0.013273557648062706
step: 60, loss: 0.0007136431522667408
step: 70, loss: 0.0023452546447515488
step: 80, loss: 0.0029711967799812555
step: 90, loss: 0.0016798536526039243
step: 100, loss: 0.0008075948571786284
step: 110, loss: 0.0006695532356388867
step: 120, loss: 0.0005719383480027318
step: 130, loss: 0.0009404059965163469
step: 140, loss: 0.0005691671976819634
step: 150, loss: 0.0034500265028327703
step: 160, loss: 0.006527693476527929
step: 170, loss: 0.00044123452971689403
step: 180, loss: 0.007753495126962662
step: 190, loss: 0.0004558579239528626
/usr/local/lib/python3.7/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=0.8220551378446115, f1=0.8268733850129197,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01683218404650688
step: 10, loss: 0.005190021358430386
step: 20, loss: 0.0026653229724615812
step: 30, loss: 0.0028121217619627714
step: 40, loss: 0.13437363505363464
step: 50, loss: 0.0021849172189831734
step: 60, loss: 0.004469145089387894
step: 70, loss: 0.011639473028481007
step: 80, loss: 0.0034550693817436695
step: 90, loss: 0.0017145427409559488
step: 100, loss: 0.0004925543325953186
step: 110, loss: 0.0005621691234409809
step: 120, loss: 0.00030418718233704567
step: 130, loss: 0.001994452904909849
step: 140, loss: 0.0007037878385744989
step: 150, loss: 0.0015507498756051064
step: 160, loss: 0.001450269017368555
step: 170, loss: 0.006427646614611149
step: 180, loss: 0.000281708111288026
step: 190, loss: 0.0005336214671842754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8232189973614776, f1=0.8225806451612903,
best_f1=0.8225806451612903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12575720250606537
step: 10, loss: 0.00977252796292305
step: 20, loss: 0.0007450944394804537
step: 30, loss: 0.00042514834785833955
step: 40, loss: 0.003336763009428978
step: 50, loss: 0.007325252518057823
step: 60, loss: 0.0003168207185808569
step: 70, loss: 0.0018326893914490938
step: 80, loss: 0.007357196416705847
step: 90, loss: 0.0004697884724009782

```

```

step: 100, loss: 0.0002889484749175608
step: 110, loss: 0.10311265289783478
step: 120, loss: 0.00530146574601531
step: 130, loss: 0.0006369291222654283
step: 140, loss: 0.003978284075856209
step: 150, loss: 0.0010082186199724674
step: 160, loss: 0.0019974808674305677
step: 170, loss: 0.0002974255185108632
step: 180, loss: 0.002197624184191227
step: 190, loss: 0.006495912093669176
/usr/local/lib/python3.7/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=0.8170426065162907, f1=0.8244274809160306,
best_f1=0.8225806451612903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014172054361552
step: 10, loss: 0.0006172092980705202
step: 20, loss: 0.008585640229284763
step: 30, loss: 0.0002887783048208803
step: 40, loss: 0.0019087096443399787
step: 50, loss: 0.0038958063814789057
step: 60, loss: 0.06814718246459961
step: 70, loss: 0.0006653208401985466
step: 80, loss: 0.0007616403745487332
step: 90, loss: 0.024713801220059395
step: 100, loss: 0.02155185304582119
step: 110, loss: 0.004751820582896471
step: 120, loss: 0.004066836088895798
step: 130, loss: 0.001425496069714427
step: 140, loss: 0.0010580718517303467
step: 150, loss: 0.0007412235136143863
step: 160, loss: 0.0034737312234938145
step: 170, loss: 0.004747394006699324
step: 180, loss: 0.001640305039472878
step: 190, loss: 0.0016292778309434652
/usr/local/lib/python3.7/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=0.82, f1=0.8277634961439588, best_f1=0.8225806451612903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 0, loss: 0.020597603172063828
step: 10, loss: 0.0006562142516486347

```

```

step: 20, loss: 0.000988139072433114
step: 30, loss: 0.0012024263851344585
step: 40, loss: 0.0010819550370797515
step: 50, loss: 0.006635343190282583
step: 60, loss: 0.0017201091395691037
step: 70, loss: 0.01658230647444725
step: 80, loss: 0.0019102232763543725
step: 90, loss: 0.0009635664755478501
step: 100, loss: 0.002199148293584585
step: 110, loss: 0.0009947266662493348
step: 120, loss: 0.0014062848640605807
step: 130, loss: 0.0003057630965486169
step: 140, loss: 0.06728531420230865
step: 150, loss: 0.0012043993920087814
step: 160, loss: 0.0006678823265247047
step: 170, loss: 0.0006675365730188787
step: 180, loss: 0.1109585240483284
step: 190, loss: 0.004308871924877167
/usr/local/lib/python3.7/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=0.821705426356589, f1=0.8368421052631578,
best_f1=0.8225806451612903
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, 220.63it/s]
load_f1 = 0.6869158878504674
real_f1 = 0.671201814058957
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 239.27it/s]
```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

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

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

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

```
--checkpoint_path checkpoints/
```

Selected optimization level 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: 0.6226938962936401
step: 10, loss: 0.3730393648147583
step: 20, loss: 0.2939871847629547
step: 30, loss: 0.36454057693481445
step: 40, loss: 0.29004719853401184
step: 50, loss: 0.2593650221824646
step: 60, loss: 0.2844364643096924
step: 70, loss: 0.3660948574542999
step: 80, loss: 0.3295149803161621
```

```

step: 90, loss: 0.2078460454940796
step: 100, loss: 0.2068057507276535
step: 110, loss: 0.22672636806964874
step: 120, loss: 0.12173006683588028
step: 130, loss: 0.024362657219171524
step: 140, loss: 0.20114260911941528
step: 150, loss: 0.13315825164318085
step: 160, loss: 0.1236371397972107
step: 170, loss: 0.18772278726100922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7919799498746868, f1=0.7648456057007125,
best_f1=0.7648456057007125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05294169485569
step: 10, loss: 0.1703510284423828
step: 20, loss: 0.059812355786561966
step: 30, loss: 0.17199759185314178
step: 40, loss: 0.0492904931306839
step: 50, loss: 0.10320239514112473
step: 60, loss: 0.07962938398122787
step: 70, loss: 0.05678071826696396
step: 80, loss: 0.024759985506534576
step: 90, loss: 0.05581369623541832
step: 100, loss: 0.07142511010169983
step: 110, loss: 0.051976121962070465
step: 120, loss: 0.09089818596839905
step: 130, loss: 0.1437087208032608
step: 140, loss: 0.3150317370891571
step: 150, loss: 0.17016619443893433
step: 160, loss: 0.10363893210887909
step: 170, loss: 0.07306842505931854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8190045248868778, f1=0.796573875802998,
best_f1=0.796573875802998

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04684998840093613
step: 10, loss: 0.041540928184986115
step: 20, loss: 0.04169442504644394
step: 30, loss: 0.11972648650407791
step: 40, loss: 0.08476997166872025
step: 50, loss: 0.04703066498041153
step: 60, loss: 0.08124809712171555
step: 70, loss: 0.03699271008372307
step: 80, loss: 0.032933760434389114
step: 90, loss: 0.06052215024828911
step: 100, loss: 0.00502611231058836
step: 110, loss: 0.055568989366292953
step: 120, loss: 0.03028569184243679
step: 130, loss: 0.024694915860891342
step: 140, loss: 0.09960253536701202
step: 150, loss: 0.004523203242570162
step: 160, loss: 0.052004940807819366
step: 170, loss: 0.13222955167293549
/usr/local/lib/python3.7/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=0.8044077134986227, f1=0.8186528497409326,
best_f1=0.796573875802998
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003846348263323307
step: 10, loss: 0.0397951602935791
step: 20, loss: 0.0019257153617218137
step: 30, loss: 0.036424413323402405
step: 40, loss: 0.0009956654394045472
step: 50, loss: 0.18895401060581207
step: 60, loss: 0.09122707694768906
step: 70, loss: 0.0016227711457759142
step: 80, loss: 0.12188717722892761
step: 90, loss: 0.12628339231014252
step: 100, loss: 0.1819351315498352
step: 110, loss: 0.07061309367418289
step: 120, loss: 0.0045013707131147385
step: 130, loss: 0.009226005524396896
step: 140, loss: 0.020462730899453163
step: 150, loss: 0.18303638696670532
step: 160, loss: 0.02416970580816269
step: 170, loss: 0.01694677770137787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8356807511737089, f1=0.8243243243243242,
best_f1=0.8243243243243242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08172105252742767
step: 10, loss: 0.05529264733195305
step: 20, loss: 0.019775236025452614
step: 30, loss: 0.0069638825953006744
step: 40, loss: 0.0009562033810652792
step: 50, loss: 0.038200508803129196
step: 60, loss: 0.020333075895905495
step: 70, loss: 0.2424611747264862
step: 80, loss: 0.0349125936627388
step: 90, loss: 0.06907419115304947
step: 100, loss: 0.02762700244784355
step: 110, loss: 0.007388115394860506
step: 120, loss: 0.0017388375708833337
step: 130, loss: 0.12102945148944855
step: 140, loss: 0.005997147876769304
step: 150, loss: 0.058888014405965805
step: 160, loss: 0.07355830073356628
step: 170, loss: 0.024564538151025772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8618266978922716, f1=0.8513513513513513,
best_f1=0.8513513513513513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009346663020551205
step: 10, loss: 0.028798799961805344
step: 20, loss: 0.02505510486662388
step: 30, loss: 0.04466408118605614
step: 40, loss: 0.01695476472377777
step: 50, loss: 0.09547529369592667
step: 60, loss: 0.027003049850463867

```



```

step: 70, loss: 0.021654140204191208
step: 80, loss: 0.004902741871774197
step: 90, loss: 0.0014872748870402575
step: 100, loss: 0.014419535174965858
step: 110, loss: 0.008263266645371914
step: 120, loss: 0.003637221409007907
step: 130, loss: 0.006654126103967428
step: 140, loss: 0.007447872310876846
step: 150, loss: 0.01234903559088707
step: 160, loss: 0.1401209831237793
step: 170, loss: 0.0019733367953449488
/usr/local/lib/python3.7/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=0.8337468982630273, f1=0.802784222737819,
best_f1=0.8513513513513513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002312253462150693
step: 10, loss: 0.00883490964770317
step: 20, loss: 0.006791639141738415
step: 30, loss: 0.004414373077452183
step: 40, loss: 0.006842443253844976
step: 50, loss: 0.013672081753611565
step: 60, loss: 0.002793190535157919
step: 70, loss: 0.0008717644959688187
step: 80, loss: 0.03248496353626251
step: 90, loss: 0.0006372914067469537
step: 100, loss: 0.0012469288194552064
step: 110, loss: 0.002224525436758995
step: 120, loss: 0.0013496438041329384
step: 130, loss: 0.0865168496966362
step: 140, loss: 0.05187615379691124
step: 150, loss: 0.002754651242867112
step: 160, loss: 0.0001766723144100979
step: 170, loss: 0.12337273359298706
/usr/local/lib/python3.7/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=0.8489208633093525, f1=0.8211009174311925,
best_f1=0.8513513513513513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021341687068343163
step: 10, loss: 0.000446033343905583
step: 20, loss: 0.0002170984080294147

```

```

step: 30, loss: 0.015533637255430222
step: 40, loss: 0.00016029247490223497
step: 50, loss: 0.0014730984112247825
step: 60, loss: 0.0006413999362848699
step: 70, loss: 0.00035229354398325086
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.01646895334124565
step: 90, loss: 0.0012300732778385282
step: 100, loss: 0.06376318633556366
step: 110, loss: 0.06845653802156448
step: 120, loss: 0.00046322689740918577
step: 130, loss: 0.016608942300081253
step: 140, loss: 0.0013594069750979543
step: 150, loss: 0.0066222213208675385
step: 160, loss: 0.01875131204724312
step: 170, loss: 0.00024083213065750897
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8627450980392157, f1=0.850467289719626,
best_f1=0.850467289719626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003779444959945977
step: 10, loss: 0.005559077952057123
step: 20, loss: 0.0009739786037243903
step: 30, loss: 0.003761737607419491
step: 40, loss: 0.03378985822200775
step: 50, loss: 0.0006999758188612759
step: 60, loss: 0.027984775602817535
step: 70, loss: 0.0014828001148998737
step: 80, loss: 0.003212706418707967
step: 90, loss: 0.028550123795866966
step: 100, loss: 0.008376708254218102
step: 110, loss: 0.0010683084838092327
step: 120, loss: 0.010541376657783985
step: 130, loss: 0.04962049424648285
step: 140, loss: 0.0035782100167125463
step: 150, loss: 0.0009509905939921737
step: 160, loss: 0.0017514722421765327
step: 170, loss: 0.014939947053790092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8746928746928747, f1=0.8530805687203792,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09117763489484787
step: 10, loss: 0.001285767531953752
step: 20, loss: 0.08201482892036438
step: 30, loss: 0.0022204394917935133
step: 40, loss: 0.0010526009136810899
step: 50, loss: 0.026007089763879776
step: 60, loss: 0.00026278485893271863
step: 70, loss: 0.01298610307276249
step: 80, loss: 0.0025171181187033653
step: 90, loss: 0.05035192146897316
step: 100, loss: 0.00017379062774125487
step: 110, loss: 0.0008166981278918684
step: 120, loss: 0.0003117911983281374
step: 130, loss: 0.00021845220180694014
step: 140, loss: 0.0033271308057010174
step: 150, loss: 0.06074291467666626
step: 160, loss: 0.03344786539673805
step: 170, loss: 0.00018941749294754118
/usr/local/lib/python3.7/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=0.8467153284671532, f1=0.8430913348946136,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024703318253159523
step: 10, loss: 0.000646770466119051
step: 20, loss: 0.0002296631719218567
step: 30, loss: 0.00020575257076416165
step: 40, loss: 0.0011894695926457644
step: 50, loss: 0.007071770261973143
step: 60, loss: 0.0028247812297195196
step: 70, loss: 0.0033906567841768265
step: 80, loss: 0.0013151172315701842
step: 90, loss: 0.06019710749387741
step: 100, loss: 0.0002192188985645771
step: 110, loss: 0.025482073426246643

```

```

step: 120, loss: 0.00024682635557837784
step: 130, loss: 0.00010985172411892563
step: 140, loss: 0.02418447658419609
step: 150, loss: 0.003857178147882223
step: 160, loss: 0.0013379368465393782
step: 170, loss: 0.0010359372245147824
/usr/local/lib/python3.7/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=0.8578431372549019, f1=0.8368794326241136,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026195228565484285
step: 10, loss: 0.00023075474018696696
step: 20, loss: 0.009683603420853615
step: 30, loss: 0.0007022919016890228
step: 40, loss: 0.00028724377625621855
step: 50, loss: 0.00019119569333270192
step: 60, loss: 0.0029077460058033466
step: 70, loss: 0.04699225723743439
step: 80, loss: 0.0002800042275339365
step: 90, loss: 0.03090381808578968
step: 100, loss: 0.00027818462694995105
step: 110, loss: 0.00022555608302354813
step: 120, loss: 0.004705336876213551
step: 130, loss: 0.0039961207658052444
step: 140, loss: 0.006311147939413786
step: 150, loss: 0.03359280526638031
step: 160, loss: 0.0002027867012657225
step: 170, loss: 0.00047611765330657363
/usr/local/lib/python3.7/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=0.8592233009708737, f1=0.8403755868544601,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02321985736489296
step: 10, loss: 0.0016021552728489041
step: 20, loss: 0.0031469985842704773
step: 30, loss: 0.00011555109813343734
step: 40, loss: 0.000722582801245153
step: 50, loss: 0.0007130313897505403
step: 60, loss: 0.005359890405088663
step: 70, loss: 0.00044059546780772507

```

```

step: 80, loss: 0.0020113475620746613
step: 90, loss: 0.00025424500927329063
step: 100, loss: 0.003065624739974737
step: 110, loss: 0.00010900716006290168
step: 120, loss: 0.03148406371474266
step: 130, loss: 0.00038786910590715706
step: 140, loss: 0.0002599587896838784
step: 150, loss: 0.00025656120851635933
step: 160, loss: 0.003049788996577263
step: 170, loss: 0.0452515184879303
/usr/local/lib/python3.7/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=0.8626506024096385, f1=0.8391608391608392,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014923898270353675
step: 10, loss: 8.398521458730102e-05
step: 20, loss: 0.007671830244362354
step: 30, loss: 0.00026758151943795383
step: 40, loss: 0.00011892959446413442
step: 50, loss: 0.00022476812591776252
step: 60, loss: 0.00021187802485655993
step: 70, loss: 0.0003515139105729759
step: 80, loss: 0.006925573572516441
step: 90, loss: 0.00014816310431342572
step: 100, loss: 0.0010311927180737257
step: 110, loss: 0.00011559748236322775
step: 120, loss: 0.059687890112400055
step: 130, loss: 0.0043736728839576244
step: 140, loss: 0.0002686023944988847
step: 150, loss: 9.591376147000119e-05
step: 160, loss: 8.232328400481492e-05
step: 170, loss: 0.0009218771010637283
/usr/local/lib/python3.7/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=0.8619854721549638, f1=0.8450704225352113,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031422129832208157
step: 10, loss: 0.000601290084887296
step: 20, loss: 0.0043101124465465546
step: 30, loss: 0.00028226894210092723

```

```

step: 40, loss: 0.00021184037905186415
step: 50, loss: 7.959151844261214e-05
step: 60, loss: 0.0050274343229830265
step: 70, loss: 7.513356831623241e-05
step: 80, loss: 0.016412898898124695
step: 90, loss: 0.007473657373338938
step: 100, loss: 0.015554788522422314
step: 110, loss: 0.0005447727162390947
step: 120, loss: 0.00018188441754318774
step: 130, loss: 0.0003576886956579983
step: 140, loss: 0.00012924747716169804
step: 150, loss: 0.00015780542162247002
step: 160, loss: 0.000724644458387047
step: 170, loss: 0.0005418010987341404
/usr/local/lib/python3.7/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=0.8578431372549019, f1=0.8436018957345971,
best_f1=0.8530805687203792
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, 264.93it/s]
load_f1 = 0.29082774049217003
real_f1 = 0.2821997105643994
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, 237.37it/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, 514kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.55MB/s]

```

Downloading: 100% 440M/440M [00:09<00:00, 47.6MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

step: 0, loss: 0.6376854181289673
step: 10, loss: 0.6053912043571472
step: 20, loss: 0.34743085503578186
step: 30, loss: 0.10645478963851929
step: 40, loss: 0.21833065152168274
step: 50, loss: 0.04472216218709946
step: 60, loss: 0.0754769966006279
step: 70, loss: 0.021940680220723152
step: 80, loss: 0.028643224388360977
step: 90, loss: 0.10485832393169403
step: 100, loss: 0.014282528311014175


```

step: 110, loss: 0.13221503794193268
step: 120, loss: 0.010548385791480541
step: 130, loss: 0.016935141757130623
step: 140, loss: 0.0038735917769372463
step: 150, loss: 0.010501017794013023
step: 160, loss: 0.006652245298027992
step: 170, loss: 0.14868097007274628
step: 180, loss: 0.11623193323612213
step: 190, loss: 0.03528976067900658
step: 200, loss: 0.08536255359649658
step: 210, loss: 0.003011586144566536
step: 220, loss: 0.00923354271799326
step: 230, loss: 0.00609257398173213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753363228699552, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005930294748395681
step: 10, loss: 0.009693928994238377
step: 20, loss: 0.051044367253780365
step: 30, loss: 0.1990283876657486
step: 40, loss: 0.006983155384659767
step: 50, loss: 0.005659614223986864
step: 60, loss: 0.004904738161712885
step: 70, loss: 0.009195569902658463
step: 80, loss: 0.002895750803872943
step: 90, loss: 0.0014992357464507222
step: 100, loss: 0.12416554242372513
step: 110, loss: 0.043830111622810364
step: 120, loss: 0.059995245188474655
step: 130, loss: 0.0010520446812734008
step: 140, loss: 0.001532653346657753
step: 150, loss: 0.005183194763958454
step: 160, loss: 0.01754535734653473
step: 170, loss: 0.0009273499017581344
step: 180, loss: 0.0032200636342167854
step: 190, loss: 0.014903008006513119
step: 200, loss: 0.0025739867705851793
step: 210, loss: 0.0009648996056057513
step: 220, loss: 0.04429788514971733

```

```

step: 230, loss: 0.002955907490104437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787234042553192, f1=0.970917225950783,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003536463947966695
step: 10, loss: 0.0021705536637455225
step: 20, loss: 0.09081707149744034
step: 30, loss: 0.023300858214497566
step: 40, loss: 0.10699491202831268
step: 50, loss: 0.0036420351825654507
step: 60, loss: 0.0013191521866247058
step: 70, loss: 0.0017223372124135494
step: 80, loss: 0.008091850206255913
step: 90, loss: 0.002648185472935438
step: 100, loss: 0.0015596410958096385
step: 110, loss: 0.0006161192432045937
step: 120, loss: 0.0705491304397583
step: 130, loss: 0.0007258729310706258
step: 140, loss: 0.007512166164815426
step: 150, loss: 0.09209680557250977
step: 160, loss: 0.06227995082736015
step: 170, loss: 0.00864116195589304
step: 180, loss: 0.009788360446691513
step: 190, loss: 0.0018748223083093762
step: 200, loss: 0.0005449085729196668
step: 210, loss: 0.01834615133702755
step: 220, loss: 0.0016404001507908106
step: 230, loss: 0.003784617641940713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9788182831661093, f1=0.9753914988814317,
best_f1=0.9753914988814317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009705494740046561
step: 10, loss: 0.0022023767232894897
step: 20, loss: 0.0011494260979816318
step: 30, loss: 0.0011779704364016652
step: 40, loss: 0.009459367021918297
step: 50, loss: 0.00042094316449947655
step: 60, loss: 0.018559208139777184
step: 70, loss: 0.0004920790088362992
step: 80, loss: 0.0013063295045867562
step: 90, loss: 0.019443197175860405
step: 100, loss: 0.0010542134987190366
step: 110, loss: 0.00027776157367043197
step: 120, loss: 0.0018681059591472149
step: 130, loss: 0.0005159444990567863
step: 140, loss: 0.0015226621180772781
step: 150, loss: 0.1099669486284256
step: 160, loss: 0.007452613208442926
step: 170, loss: 0.002931026043370366
step: 180, loss: 0.00043835723772644997
step: 190, loss: 0.0009547676891088486
step: 200, loss: 0.00031441583996638656
step: 210, loss: 0.04202418029308319
step: 220, loss: 0.0003011936496477574
step: 230, loss: 0.00351236411370337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9797752808988766, f1=0.9752808988764046,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009912971872836351
step: 10, loss: 0.0037656344939023256
step: 20, loss: 0.002044856548309326
step: 30, loss: 0.0012740641832351685
step: 40, loss: 0.0072714174166321754
step: 50, loss: 0.0010742812883108854
step: 60, loss: 0.016169780865311623
step: 70, loss: 0.0007554598851129413
step: 80, loss: 0.007588118780404329
step: 90, loss: 0.0024210938718169928
step: 100, loss: 0.00027405930450186133

```

```

step: 110, loss: 0.0077769141644239426
step: 120, loss: 0.0002728867984842509
step: 130, loss: 0.00039974020910449326
step: 140, loss: 0.00044606681331060827
step: 150, loss: 0.010088622570037842
step: 160, loss: 0.00168724509421736
step: 170, loss: 0.012091251090168953
step: 180, loss: 0.00014069673488847911
step: 190, loss: 0.10918352752923965
step: 200, loss: 0.008008862845599651
step: 210, loss: 0.00064947857754305
step: 220, loss: 0.0020187075715512037
step: 230, loss: 0.0002870619937311858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.983277591973244, f1=0.9821029082774049,
best_f1=0.9821029082774049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004578636959195137
step: 10, loss: 0.00025681592524051666
step: 20, loss: 0.00389959872700274
step: 30, loss: 0.00012217293260619044
step: 40, loss: 0.00021612834825646132
step: 50, loss: 0.0007390903774648905
step: 60, loss: 0.00023601087741553783
step: 70, loss: 0.00025152810849249363
step: 80, loss: 0.00032352301059290767
step: 90, loss: 0.00024010761990211904
step: 100, loss: 0.16285964846611023
step: 110, loss: 0.0003219636273570359
step: 120, loss: 0.0008074999204836786
step: 130, loss: 0.0016116249607875943
step: 140, loss: 0.00020355671586003155
step: 150, loss: 0.15025268495082855
step: 160, loss: 0.0028257460799068213
step: 170, loss: 0.0002701892517507076
step: 180, loss: 0.14538033306598663
step: 190, loss: 0.04919467493891716
step: 200, loss: 0.0006717260112054646
step: 210, loss: 0.0006698522483929992
step: 220, loss: 0.0013711736537516117

```

```

step: 230, loss: 0.04831946641206741
/usr/local/lib/python3.7/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=0.9830890642615557, f1=0.9796380090497738,
best_f1=0.9821029082774049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019055861048400402
step: 10, loss: 0.00025552502484060824
step: 20, loss: 0.0014513246715068817
step: 30, loss: 0.001885338337160647
step: 40, loss: 0.0009334494243375957
step: 50, loss: 0.0005564873572438955
step: 60, loss: 0.0008947290480136871
step: 70, loss: 0.013261542655527592
step: 80, loss: 0.0052832188084721565
step: 90, loss: 0.000609140784945339
step: 100, loss: 0.00012430209608282894
step: 110, loss: 0.0028152884915471077
step: 120, loss: 8.709971734788269e-05
step: 130, loss: 0.00011474839993752539
step: 140, loss: 0.0029467330314219
step: 150, loss: 0.0039027617312967777
step: 160, loss: 0.076616570353508
step: 170, loss: 0.010296177119016647
step: 180, loss: 0.00018323837139178067
step: 190, loss: 7.400499453069642e-05
step: 200, loss: 0.038978226482868195
step: 210, loss: 6.288549047894776e-05
step: 220, loss: 0.00012884761963505298
step: 230, loss: 0.0003399307606741786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9876265466816648, f1=0.9809203142536477,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.905836926307529e-05
step: 10, loss: 0.0004365033528301865
step: 20, loss: 0.0005265409126877785

```

```

step: 30, loss: 0.00011409179569454864
step: 40, loss: 0.00039105876930989325
step: 50, loss: 0.0001283948076888919
step: 60, loss: 5.9618716477416456e-05
step: 70, loss: 0.00011641927994787693
step: 80, loss: 0.0001328616781393066
step: 90, loss: 5.1913666538894176e-05
step: 100, loss: 6.465410842793062e-05
step: 110, loss: 0.0069536371156573296
step: 120, loss: 0.004570046439766884
step: 130, loss: 0.000467084493720904
step: 140, loss: 5.2232549933250993e-05
step: 150, loss: 5.6698339903960004e-05
step: 160, loss: 8.830225124256685e-05
step: 170, loss: 0.0002509288606233895
step: 180, loss: 0.013523942790925503
step: 190, loss: 0.0005201388848945498
step: 200, loss: 0.00017387114348821342
step: 210, loss: 0.007365836761891842
step: 220, loss: 8.312166028190404e-05
step: 230, loss: 7.36331712687388e-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=0.9864864864864865, f1=0.9786276715410572,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.4860124287661165e-05
step: 10, loss: 0.0006590400589630008
step: 20, loss: 0.00019275155500508845
step: 30, loss: 4.860846820520237e-05
step: 40, loss: 0.04124198853969574
step: 50, loss: 7.715663377894089e-05
step: 60, loss: 0.00010914399899775162
step: 70, loss: 0.00012046264600940049
step: 80, loss: 6.350412149913609e-05
step: 90, loss: 0.004364732652902603
step: 100, loss: 0.00011179613647982478
step: 110, loss: 0.0005202332977205515
step: 120, loss: 3.9777278288966045e-05
step: 130, loss: 0.00017820618813857436
step: 140, loss: 3.579878466553055e-05
step: 150, loss: 0.03307710215449333
step: 160, loss: 0.00011231208918616176
step: 170, loss: 0.0006728385342285037
step: 180, loss: 0.00013116335321683437

```

```

step: 190, loss: 5.879166928934865e-05
step: 200, loss: 0.0007175736245699227
step: 210, loss: 0.0007789563969708979
step: 220, loss: 5.844104089192115e-05
step: 230, loss: 0.00061297498177737
/usr/local/lib/python3.7/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=0.9831649831649831, f1=0.9809203142536477,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003912633692380041
step: 10, loss: 6.8330904468894e-05
step: 20, loss: 0.0002107359905494377
step: 30, loss: 5.0715418183244765e-05
step: 40, loss: 6.594417936867103e-05
step: 50, loss: 0.0001674556260695681
step: 60, loss: 0.00044793059350922704
step: 70, loss: 0.00032293799449689686
step: 80, loss: 4.291369259590283e-05
step: 90, loss: 7.637766975676641e-05
step: 100, loss: 8.341654029209167e-05
step: 110, loss: 5.359042916097678e-05
step: 120, loss: 0.0001878520124591887
step: 130, loss: 4.130031811655499e-05
step: 140, loss: 0.03580639511346817
step: 150, loss: 0.013157851994037628
step: 160, loss: 0.00023932851036079228
step: 170, loss: 6.243935786187649e-05
step: 180, loss: 5.832695751450956e-05
step: 190, loss: 0.007009926252067089
step: 200, loss: 3.5541590477805585e-05
step: 210, loss: 0.00012167319073341787
step: 220, loss: 3.8975751522229984e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 230, loss: 5.245109059615061e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9864864864864865, f1=0.9809203142536477,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.534037386998534e-05
step: 10, loss: 4.960345177096315e-05

```

```

step: 20, loss: 0.00011456905485829338
step: 30, loss: 0.00022084206284489483
step: 40, loss: 0.00015952701505739242
step: 50, loss: 8.065909787546843e-05
step: 60, loss: 0.014546790160238743
step: 70, loss: 4.097568671568297e-05
step: 80, loss: 3.8326634239638224e-05
step: 90, loss: 0.00015194117440842092
step: 100, loss: 4.220938353682868e-05
step: 110, loss: 0.005333541426807642
step: 120, loss: 5.2547937229974195e-05
step: 130, loss: 4.679893027059734e-05
step: 140, loss: 3.59737932740245e-05
step: 150, loss: 0.02147945761680603
step: 160, loss: 6.488532380899414e-05
step: 170, loss: 0.019178234040737152
step: 180, loss: 7.064088276820257e-05
step: 190, loss: 3.341487536090426e-05
step: 200, loss: 0.0005504721193574369
step: 210, loss: 6.994641444180161e-05
step: 220, loss: 3.537792872521095e-05
step: 230, loss: 2.865041642508004e-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=0.9864864864864865, f1=0.9763779527559054,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.291903678677045e-05
step: 10, loss: 2.9909531804150902e-05
step: 20, loss: 4.489077764446847e-05
step: 30, loss: 4.838101813220419e-05
step: 40, loss: 0.00018922802701126784
step: 50, loss: 0.0010358849540352821
step: 60, loss: 0.00021296579507179558
step: 70, loss: 4.0280214307131246e-05
step: 80, loss: 7.461931818397716e-05
step: 90, loss: 2.9760520192212425e-05
step: 100, loss: 4.825424912269227e-05
step: 110, loss: 4.326487396610901e-05
step: 120, loss: 3.453559475019574e-05
step: 130, loss: 3.285979983047582e-05
step: 140, loss: 0.01730486750602722
step: 150, loss: 6.729571032337844e-05
step: 160, loss: 3.523953273543157e-05
step: 170, loss: 4.073388481629081e-05

```



```

step: 180, loss: 0.01904415711760521
step: 190, loss: 0.00028170953737571836
step: 200, loss: 2.7171596229891293e-05
step: 210, loss: 0.00016290944768115878
step: 220, loss: 2.8575959731824696e-05
step: 230, loss: 0.04187273979187012
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987709497206704, f1=0.9788182831661093,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010297949484083802
step: 10, loss: 4.7548288421239704e-05
step: 20, loss: 0.00030827909358777106
step: 30, loss: 4.861613706452772e-05
step: 40, loss: 3.2874417229322717e-05
step: 50, loss: 0.00017433243920095265
step: 60, loss: 6.562700582435355e-05
step: 70, loss: 4.670289490604773e-05
step: 80, loss: 0.00021956766431685537
step: 90, loss: 3.806417225860059e-05
step: 100, loss: 2.367737761233002e-05
step: 110, loss: 0.027727793902158737
step: 120, loss: 0.015647532418370247
step: 130, loss: 5.895634967600927e-05
step: 140, loss: 3.497907891869545e-05
step: 150, loss: 3.0978677386883646e-05
step: 160, loss: 3.07656591758132e-05
step: 170, loss: 0.0001069301288225688
step: 180, loss: 3.7261503166519105e-05
step: 190, loss: 3.370140984770842e-05
step: 200, loss: 0.0001118296422646381
step: 210, loss: 2.5621980967116542e-05
step: 220, loss: 3.5673772799782455e-05
step: 230, loss: 5.592902016360313e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9854096520763187, f1=0.9776286353467561,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.8322614525677636e-05
step: 10, loss: 9.792469791136682e-05
step: 20, loss: 3.933200059691444e-05
step: 30, loss: 5.992744991090149e-05
step: 40, loss: 0.01516694575548172
step: 50, loss: 7.825871580280364e-05
step: 60, loss: 4.787493890034966e-05
step: 70, loss: 3.3816497307270765e-05
step: 80, loss: 3.272193134762347e-05
step: 90, loss: 3.819345511146821e-05
step: 100, loss: 2.5815621484071016e-05
step: 110, loss: 0.00012361312110442668
step: 120, loss: 2.0443992980290204e-05
step: 130, loss: 2.9537128284573555e-05
step: 140, loss: 2.8613238100660965e-05
step: 150, loss: 2.6974117645295337e-05
step: 160, loss: 0.0004812393453903496
step: 170, loss: 2.704053622437641e-05
step: 180, loss: 3.3019798138411716e-05
step: 190, loss: 0.00010677776299417019
step: 200, loss: 2.3848642740631476e-05
step: 210, loss: 0.00019230027101002634
step: 220, loss: 0.00018183220527134836
step: 230, loss: 0.008472444489598274
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9785794813979707,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.9278062609373592e-05
step: 10, loss: 1.6953566955635324e-05
step: 20, loss: 2.9302153052412905e-05
step: 30, loss: 4.036473546875641e-05
step: 40, loss: 4.609623283613473e-05
step: 50, loss: 0.02984575927257538
step: 60, loss: 4.449281914276071e-05
step: 70, loss: 3.112329068244435e-05
step: 80, loss: 4.785326382261701e-05
step: 90, loss: 0.0001565888524055481
step: 100, loss: 3.651672159321606e-05
step: 110, loss: 2.3737035007798113e-05
step: 120, loss: 3.875941183650866e-05
step: 130, loss: 0.011057173833251
```

```

step: 140, loss: 2.9283331969054416e-05
step: 150, loss: 5.195193807594478e-05
step: 160, loss: 1.987404903047718e-05
step: 170, loss: 3.3055097446776927e-05
step: 180, loss: 0.0001374236453557387
step: 190, loss: 4.111956513952464e-05
step: 200, loss: 7.759435538901016e-05
step: 210, loss: 3.7161007639952004e-05
step: 220, loss: 0.00024253144511021674
step: 230, loss: 2.282809691678267e-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=0.987598647125141, f1=0.9785794813979707,
best_f1=0.9788182831661093
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 193.04it/s]
load_f1 = 0.9876265466816648
real_f1 = 0.9876265466816648

```

```

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:19, 222.47it/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: 0.6223543882369995
step: 10, loss: 0.5245537161827087
step: 20, loss: 0.4439574182033539
step: 30, loss: 0.12671849131584167
step: 40, loss: 0.18743403255939484
step: 50, loss: 0.15721970796585083
step: 60, loss: 0.026217443868517876
step: 70, loss: 0.08980930596590042
step: 80, loss: 0.026548193767666817
step: 90, loss: 0.1171327531337738
step: 100, loss: 0.015269520692527294
step: 110, loss: 0.05860515683889389
step: 120, loss: 0.10557756572961807
step: 130, loss: 0.12499988079071045
step: 140, loss: 0.06719816476106644
step: 150, loss: 0.029660260304808617
step: 160, loss: 0.09189265221357346
step: 170, loss: 0.16143278777599335
step: 180, loss: 0.048619598150253296
step: 190, loss: 0.04322631657123566

```

```

step: 200, loss: 0.11057496070861816
step: 210, loss: 0.058633480221033096
step: 220, loss: 0.17666637897491455
step: 230, loss: 0.13676811754703522
step: 240, loss: 0.06302032619714737
step: 250, loss: 0.013514453545212746
step: 260, loss: 0.02475108578801155
step: 270, loss: 0.03951451554894447
step: 280, loss: 0.05772310122847557
step: 290, loss: 0.03495687618851662
step: 300, loss: 0.0413409098982811
step: 310, loss: 0.07254505902528763
step: 320, loss: 0.11396907269954681
step: 330, loss: 0.013665479607880116
step: 340, loss: 0.016750672832131386
step: 350, loss: 0.0491425059735775
step: 360, loss: 0.02332419343292713
step: 370, loss: 0.19790112972259521
step: 380, loss: 0.017518136650323868
step: 390, loss: 0.10429708659648895
step: 400, loss: 0.2242615669965744
step: 410, loss: 0.03211863338947296
step: 420, loss: 0.008830866776406765
step: 430, loss: 0.15880268812179565
step: 440, loss: 0.03275444731116295
step: 450, loss: 0.006043011322617531
step: 460, loss: 0.009471509605646133
step: 470, loss: 0.11279474198818207
step: 480, loss: 0.0404355488717556
step: 490, loss: 0.056404534727334976
step: 500, loss: 0.07149584591388702
step: 510, loss: 0.042825158685445786
step: 520, loss: 0.10390293598175049
step: 530, loss: 0.011421342380344868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9434482758620689, f1=0.9480937069361507,
best_f1=0.9480937069361507
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.055374626070261
step: 10, loss: 0.04458664730191231

```

step: 20, loss: 0.0406232587993145
step: 30, loss: 0.06462309509515762
step: 40, loss: 0.06313405930995941
step: 50, loss: 0.15684296190738678
step: 60, loss: 0.006811518222093582
step: 70, loss: 0.040493905544281006
step: 80, loss: 0.060683153569698334
step: 90, loss: 0.005719384644180536
step: 100, loss: 0.02133970707654953
step: 110, loss: 0.005425148643553257
step: 120, loss: 0.06797680258750916
step: 130, loss: 0.06402605772018433
step: 140, loss: 0.008782004937529564
step: 150, loss: 0.048713065683841705
step: 160, loss: 0.007027555722743273
step: 170, loss: 0.005193960852921009
step: 180, loss: 0.024412555620074272
step: 190, loss: 0.06688863784074783
step: 200, loss: 0.007478930987417698
step: 210, loss: 0.02749812789261341
step: 220, loss: 0.2086547464132309
step: 230, loss: 0.010426000691950321
step: 240, loss: 0.06761865317821503
step: 250, loss: 0.05059430003166199
step: 260, loss: 0.0011266034562140703
step: 270, loss: 0.05032432824373245
step: 280, loss: 0.010491310618817806
step: 290, loss: 0.03537292033433914
step: 300, loss: 0.18957316875457764
step: 310, loss: 0.01944754086434841
step: 320, loss: 0.05365830659866333
step: 330, loss: 0.018129639327526093
step: 340, loss: 0.006701637990772724
step: 350, loss: 0.0005966636235825717
step: 360, loss: 0.03501015529036522
step: 370, loss: 0.09011615812778473
step: 380, loss: 0.044393252581357956
step: 390, loss: 0.07499440759420395
step: 400, loss: 0.022445982322096825
step: 410, loss: 0.02606194093823433
step: 420, loss: 0.08420053869485855
step: 430, loss: 0.008480310440063477
step: 440, loss: 0.1659255176782608
step: 450, loss: 0.009475313127040863
step: 460, loss: 0.0850893035531044
step: 470, loss: 0.11428991705179214
step: 480, loss: 0.2582464814186096
step: 490, loss: 0.01573205552995205

```

step: 500, loss: 0.13175831735134125
step: 510, loss: 0.0214703232049942
step: 520, loss: 0.0830642580986023
step: 530, loss: 0.03629835695028305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9472710453283996, f1=0.9458218549127642,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0487125962972641
step: 10, loss: 0.04956236109137535
step: 20, loss: 0.15291064977645874
step: 30, loss: 0.09628495573997498
step: 40, loss: 0.03311162069439888
step: 50, loss: 0.1310197412967682
step: 60, loss: 0.02857782505452633
step: 70, loss: 0.008059250190854073
step: 80, loss: 0.0009617740288376808
step: 90, loss: 0.0057777767069637775
step: 100, loss: 0.06105673313140869
step: 110, loss: 0.0038446735125035048
step: 120, loss: 0.011518999002873898
step: 130, loss: 0.002097098156809807
step: 140, loss: 0.0623137541115284
step: 150, loss: 0.02082797884941101
step: 160, loss: 0.014619342051446438
step: 170, loss: 0.07350116968154907
step: 180, loss: 0.006611284334212542
step: 190, loss: 0.008501050993800163
step: 200, loss: 0.06651279330253601
step: 210, loss: 0.06209428608417511
step: 220, loss: 0.03233868628740311
step: 230, loss: 0.09064274281263351
step: 240, loss: 0.0009183295769616961
step: 250, loss: 0.010229300707578659
step: 260, loss: 0.01728592813014984
step: 270, loss: 0.003647822653874755
step: 280, loss: 0.20634855329990387
step: 290, loss: 0.0030594058334827423
step: 300, loss: 0.009839416481554508
step: 310, loss: 0.004286954179406166

```



```

step: 320, loss: 0.015316037461161613
step: 330, loss: 0.009007211774587631
step: 340, loss: 0.005037063732743263
step: 350, loss: 0.0013216331135481596
step: 360, loss: 0.0215583648532629
step: 370, loss: 0.009588191285729408
step: 380, loss: 0.005035215988755226
step: 390, loss: 0.00523523660376668
step: 400, loss: 0.03855874389410019
step: 410, loss: 0.0038752341642975807
step: 420, loss: 0.15021921694278717
step: 430, loss: 0.005740684922784567
step: 440, loss: 0.009722334332764149
step: 450, loss: 0.09515189379453659
step: 460, loss: 0.0566793754696846
step: 470, loss: 0.06656317412853241
step: 480, loss: 0.014301951043307781
step: 490, loss: 0.014742104336619377
step: 500, loss: 0.003864399855956435
step: 510, loss: 0.006397646386176348
step: 520, loss: 0.06849581003189087
step: 530, loss: 0.03213398531079292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9492787342950209, f1=0.9471243042671614,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02520880475640297
step: 10, loss: 0.010774045251309872
step: 20, loss: 0.005340721923857927
step: 30, loss: 0.0017281241016462445
step: 40, loss: 0.004825244657695293
step: 50, loss: 0.03331034630537033
step: 60, loss: 0.11870349943637848
step: 70, loss: 0.03664957731962204
step: 80, loss: 0.006886930204927921
step: 90, loss: 0.007418026216328144
step: 100, loss: 0.010906767100095749
step: 110, loss: 0.007160353474318981
step: 120, loss: 0.00047755573177710176
step: 130, loss: 0.0021206135861575603

```

```

step: 140, loss: 0.007108731660991907
step: 150, loss: 0.013575887307524681
step: 160, loss: 0.07114998251199722
step: 170, loss: 0.0035079065710306168
step: 180, loss: 0.004756099544465542
step: 190, loss: 0.044523101300001144
step: 200, loss: 0.005749272648245096
step: 210, loss: 0.03009740076959133
step: 220, loss: 0.0021759585943073034
step: 230, loss: 0.05049527809023857
step: 240, loss: 0.0018773849587887526
step: 250, loss: 0.0037878090515732765
step: 260, loss: 0.0012461937731131911
step: 270, loss: 0.004205530043691397
step: 280, loss: 0.15078726410865784
step: 290, loss: 0.014560858719050884
step: 300, loss: 0.00037520079058595
step: 310, loss: 0.006091815419495106
step: 320, loss: 0.014322208240628242
step: 330, loss: 0.007288704160600901
step: 340, loss: 0.005224814638495445
step: 350, loss: 0.015962665900588036
step: 360, loss: 0.03482016175985336
step: 370, loss: 0.0055860234424471855
step: 380, loss: 0.05374914035201073
step: 390, loss: 0.0018307672580704093
step: 400, loss: 0.03685298562049866
step: 410, loss: 0.006981092970818281
step: 420, loss: 0.00037510247784666717
step: 430, loss: 0.05509062856435776
step: 440, loss: 0.0019861457403749228
step: 450, loss: 0.00150301493704319
step: 460, loss: 0.017518678680062294
step: 470, loss: 0.023591071367263794
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 480, loss: 0.2588529884815216
step: 490, loss: 0.01626589149236679
step: 500, loss: 0.011722886934876442
step: 510, loss: 0.018871532753109932
step: 520, loss: 0.05754786729812622
step: 530, loss: 0.011481077410280704
/usr/local/lib/python3.7/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=0.9451360073766714, f1=0.9437070938215102,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1309586465358734
step: 10, loss: 0.06216778978705406
step: 20, loss: 0.014274440705776215
step: 30, loss: 0.00796897616237402
step: 40, loss: 0.15547579526901245
step: 50, loss: 0.006701808422803879
step: 60, loss: 0.02999166026711464
step: 70, loss: 0.0019515961175784469
step: 80, loss: 0.0005441835382953286
step: 90, loss: 0.022462142631411552
step: 100, loss: 0.08622575551271439
step: 110, loss: 0.009038355201482773
step: 120, loss: 0.005413897801190615
step: 130, loss: 0.01605919376015663
step: 140, loss: 0.004370828624814749
step: 150, loss: 0.02437511831521988
step: 160, loss: 0.0004698603297583759
step: 170, loss: 0.13700374960899353
step: 180, loss: 0.005391269456595182
step: 190, loss: 0.003966202028095722
step: 200, loss: 0.0005849857116118073
step: 210, loss: 0.001849111053161323
step: 220, loss: 0.001437401631847024
step: 230, loss: 0.013971072621643543
step: 240, loss: 0.0007365496712736785
step: 250, loss: 0.00034443617914803326
step: 260, loss: 0.0036596066784113646
step: 270, loss: 0.0016279795672744513
step: 280, loss: 0.0005086626624688506
step: 290, loss: 0.20886144042015076
step: 300, loss: 0.0034041861072182655
step: 310, loss: 0.00805964507162571
step: 320, loss: 0.16646769642829895
step: 330, loss: 0.06358818709850311
step: 340, loss: 0.002160963136702776
step: 350, loss: 0.0012892583617940545
step: 360, loss: 0.0022770895157009363
step: 370, loss: 0.01061477605253458
step: 380, loss: 0.0012158015742897987
step: 390, loss: 0.0002954024530481547
step: 400, loss: 0.061776671558618546
step: 410, loss: 0.0016991081647574902
step: 420, loss: 0.0009925704216584563
step: 430, loss: 0.0032270492520183325
step: 440, loss: 0.006248191464692354
step: 450, loss: 0.07693782448768616
step: 460, loss: 0.0014036917127668858
```

```
step: 470, loss: 0.002412244211882353
step: 480, loss: 0.072774238884449
step: 490, loss: 0.0004733610257972032
step: 500, loss: 0.006232140585780144
step: 510, loss: 0.00578649528324604
step: 520, loss: 0.10620124638080597
step: 530, loss: 0.04848681017756462
/usr/local/lib/python3.7/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=0.9465364946536494, f1=0.9461966604823748,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016614908818155527
step: 10, loss: 0.0007279973360709846
step: 20, loss: 0.07771387696266174
step: 30, loss: 0.0016031223349273205
step: 40, loss: 0.08863838762044907
step: 50, loss: 0.009062924422323704
step: 60, loss: 0.0010600545210763812
step: 70, loss: 0.002782614668831229
step: 80, loss: 0.03722434863448143
step: 90, loss: 0.0003241074155084789
step: 100, loss: 0.0005483233253471553
step: 110, loss: 0.011880003847181797
step: 120, loss: 0.0015849347691982985
step: 130, loss: 0.0004080475482624024
step: 140, loss: 0.0018456658581271768
step: 150, loss: 0.0003120197798125446
step: 160, loss: 0.0014828310813754797
step: 170, loss: 0.01234754454344511
step: 180, loss: 0.007619595155119896
step: 190, loss: 0.030374327674508095
step: 200, loss: 0.0018867618637159467
step: 210, loss: 0.019305462017655373
step: 220, loss: 0.0011830549919977784
step: 230, loss: 0.000804921961389482
step: 240, loss: 0.054664891213178635
step: 250, loss: 0.002539699198678136
step: 260, loss: 0.00038976658834144473
step: 270, loss: 0.05860011279582977
step: 280, loss: 0.0028355743270367384
step: 290, loss: 0.0004171036125626415
step: 300, loss: 0.009415174834430218
step: 310, loss: 0.0008985038730315864
step: 320, loss: 0.0005066413432359695
```

```

step: 330, loss: 0.0006159063777886331
step: 340, loss: 0.1794731169939041
step: 350, loss: 0.01389908604323864
step: 360, loss: 0.005422539543360472
step: 370, loss: 0.003547669854015112
step: 380, loss: 0.002621743129566312
step: 390, loss: 0.045653678476810455
step: 400, loss: 0.00012016364780720323
step: 410, loss: 0.00015146385703701526
step: 420, loss: 0.0016571568557992578
step: 430, loss: 0.001008391729556024
step: 440, loss: 0.0019933602306991816
step: 450, loss: 0.002124006859958172
step: 460, loss: 0.0005376641056500375
step: 470, loss: 0.0008916955557651818
step: 480, loss: 0.007276324555277824
step: 490, loss: 0.007005432620644569
step: 500, loss: 0.01614212617278099
step: 510, loss: 0.006699074059724808
step: 520, loss: 0.006344974972307682
step: 530, loss: 0.006364071276038885
/usr/local/lib/python3.7/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=0.9407407407407408, f1=0.9373271889400921,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03002779744565487
step: 10, loss: 0.013320717960596085
step: 20, loss: 0.013758423738181591
step: 30, loss: 0.009453068487346172
step: 40, loss: 0.004463831894099712
step: 50, loss: 0.006905275397002697
step: 60, loss: 0.005150914192199707
step: 70, loss: 0.02758229337632656
step: 80, loss: 0.021528832614421844
step: 90, loss: 0.017303545027971268
step: 100, loss: 0.0008247719379141927
step: 110, loss: 0.00556193245574832
step: 120, loss: 0.0052943057380616665
step: 130, loss: 0.003398876404389739
step: 140, loss: 0.00011917911615455523
step: 150, loss: 0.0010128639405593276
step: 160, loss: 0.00022514436568599194
step: 170, loss: 0.00032003343221731484
step: 180, loss: 0.0025311193894594908

```

```

step: 190, loss: 0.019905325025320053
step: 200, loss: 0.0005377212655730546
step: 210, loss: 0.001315502217039466
step: 220, loss: 0.00018117058789357543
step: 230, loss: 0.0028818058781325817
step: 240, loss: 0.003211948787793517
step: 250, loss: 0.0013520527863875031
step: 260, loss: 0.0007411011029034853
step: 270, loss: 0.006348572205752134
step: 280, loss: 0.04360232129693031
step: 290, loss: 0.0005113785155117512
step: 300, loss: 0.004864044021815062
step: 310, loss: 0.0014760523336008191
step: 320, loss: 0.0005855990457348526
step: 330, loss: 0.003408692544326186
step: 340, loss: 0.017694935202598572
step: 350, loss: 0.09997362643480301
step: 360, loss: 0.0018294142792001367
step: 370, loss: 0.0009259053040295839
step: 380, loss: 0.015384732745587826
step: 390, loss: 0.019924642518162727
step: 400, loss: 0.03829782456159592
step: 410, loss: 0.007876954972743988
step: 420, loss: 0.00040955052827484906
step: 430, loss: 0.0009407186298631132
step: 440, loss: 0.0347285158932209
step: 450, loss: 0.00046547973761335015
step: 460, loss: 0.0008915679063647985
step: 470, loss: 0.1172700971364975
step: 480, loss: 0.0049169003032147884
step: 490, loss: 0.0006549095269292593
step: 500, loss: 0.000281847344012931
step: 510, loss: 0.0011346880346536636
step: 520, loss: 0.05592433363199234
step: 530, loss: 0.016163857653737068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9527374824520355, f1=0.9478584729981377,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0300087071955204

```

step: 10, loss: 0.01442454569041729
step: 20, loss: 9.173463331535459e-05
step: 30, loss: 0.0072095864452421665
step: 40, loss: 0.0002620189043227583
step: 50, loss: 0.0022188087459653616
step: 60, loss: 0.0016151079908013344
step: 70, loss: 0.0010531020816415548
step: 80, loss: 0.0051213097758591175
step: 90, loss: 0.0013289632042869925
step: 100, loss: 0.0010567536810413003
step: 110, loss: 0.0004652234201785177
step: 120, loss: 0.0012516166316345334
step: 130, loss: 0.0003523966297507286
step: 140, loss: 0.006946546956896782
step: 150, loss: 0.000544374983292073
step: 160, loss: 0.019592152908444405
step: 170, loss: 0.028969736769795418
step: 180, loss: 0.009772938676178455
step: 190, loss: 0.0032484710682183504
step: 200, loss: 0.0013635643990710378
step: 210, loss: 0.0030073777306824923
step: 220, loss: 0.002986025996506214
step: 230, loss: 0.00017424793622922152
step: 240, loss: 0.0019482895731925964
step: 250, loss: 0.00010583360563032329
step: 260, loss: 0.0005634050467051566
step: 270, loss: 0.010179603472352028
step: 280, loss: 0.005500916391611099
step: 290, loss: 0.0020160803105682135
step: 300, loss: 0.000497960252687335
step: 310, loss: 0.04949337616562843
step: 320, loss: 0.023480037227272987
step: 330, loss: 0.09548541158437729
step: 340, loss: 0.002807390643283725
step: 350, loss: 0.00028192909667268395
step: 360, loss: 0.003290773369371891
step: 370, loss: 0.012016583234071732
step: 380, loss: 0.000349964335327968
step: 390, loss: 0.16314361989498138
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 400, loss: 0.0006324333953671157
step: 410, loss: 0.0006669423310086131
step: 420, loss: 0.0021287656854838133
step: 430, loss: 0.02407647669315338
step: 440, loss: 0.002007595030590892
step: 450, loss: 0.011522595770657063
step: 460, loss: 0.0011725452495738864
step: 470, loss: 0.0021010295022279024

```

step: 480, loss: 7.512828597100452e-05
step: 490, loss: 0.13552846014499664
step: 500, loss: 0.0021315687336027622
step: 510, loss: 0.04070619121193886
step: 520, loss: 0.006643576081842184
step: 530, loss: 0.0009217982296831906
/usr/local/lib/python3.7/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=0.9451476793248945, f1=0.9479553903345725,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003625069512054324
step: 10, loss: 0.0002216337015852332
step: 20, loss: 0.000711653905455023
step: 30, loss: 0.0052890595979988575
step: 40, loss: 0.0005500013357959688
step: 50, loss: 0.00014096895756665617
step: 60, loss: 0.0008537282701581717
step: 70, loss: 0.006701187696307898
step: 80, loss: 0.002804323798045516
step: 90, loss: 0.00018225649546366185
step: 100, loss: 0.00013010308612138033
step: 110, loss: 7.422559428960085e-05
step: 120, loss: 0.0021478186827152967
step: 130, loss: 0.06232347711920738
step: 140, loss: 0.0015504250768572092
step: 150, loss: 0.0010384980123490095
step: 160, loss: 0.0015207695541903377
step: 170, loss: 0.007623632438480854
step: 180, loss: 0.0016950910212472081
step: 190, loss: 0.00039955280954018235
step: 200, loss: 0.0006820489652454853
step: 210, loss: 0.004644160624593496
step: 220, loss: 0.0029142347630113363
step: 230, loss: 0.00020761488121934235
step: 240, loss: 0.025940805673599243
step: 250, loss: 0.004571431316435337
step: 260, loss: 0.043954916298389435
step: 270, loss: 9.748184675117955e-05
step: 280, loss: 0.0006287320284172893
step: 290, loss: 0.02561267651617527
step: 300, loss: 0.00019564676040317863
step: 310, loss: 0.02366241067647934
step: 320, loss: 0.0007382696494460106
step: 330, loss: 0.0008649712544865906

```



```

step: 340, loss: 0.0019688524771481752
step: 350, loss: 0.0034051798284053802
step: 360, loss: 0.00010377925354987383
step: 370, loss: 0.0011596678523346782
step: 380, loss: 0.00013407420192379504
step: 390, loss: 0.00038639825652353466
step: 400, loss: 0.012057038024067879
step: 410, loss: 0.00011727430683095008
step: 420, loss: 0.00010580869275145233
step: 430, loss: 6.562861381098628e-05
step: 440, loss: 0.0019131343578919768
step: 450, loss: 0.0831919014453888
step: 460, loss: 0.0005561579600907862
step: 470, loss: 0.00014940091932658106
step: 480, loss: 6.35277247056365e-05
step: 490, loss: 0.04076367989182472
step: 500, loss: 0.002845807932317257
step: 510, loss: 0.005088026635348797
step: 520, loss: 0.004364265128970146
step: 530, loss: 0.0005645840428769588
/usr/local/lib/python3.7/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=0.9442139234670355, f1=0.9388505747126438,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006681061931885779
step: 10, loss: 5.054028952145018e-05
step: 20, loss: 2.8717735403915867e-05
step: 30, loss: 0.00688098045065999
step: 40, loss: 0.0003407240437809378
step: 50, loss: 0.0003758756793104112
step: 60, loss: 0.00029069127049297094
step: 70, loss: 0.0002542588044889271
step: 80, loss: 5.1569997594924644e-05
step: 90, loss: 0.0007247291505336761
step: 100, loss: 0.000271362136118114
step: 110, loss: 8.768558473093435e-05
step: 120, loss: 0.0011359883937984705
step: 130, loss: 0.001866518403403461
step: 140, loss: 0.009936334565281868
step: 150, loss: 0.006364355329424143
step: 160, loss: 6.368852336890996e-05
step: 170, loss: 0.0004562675894703716
step: 180, loss: 9.238994971383363e-05
step: 190, loss: 0.0034329185727983713

```

```

step: 200, loss: 0.00035258219577372074
step: 210, loss: 0.00015542551409453154
step: 220, loss: 0.0037748408503830433
step: 230, loss: 0.0011362220393493772
step: 240, loss: 0.00019367580534890294
step: 250, loss: 0.00020032537577208132
step: 260, loss: 0.0011956071248278022
step: 270, loss: 0.0002989580971188843
step: 280, loss: 0.050601776689291
step: 290, loss: 7.630760228494182e-05
step: 300, loss: 4.1478946513962e-05
step: 310, loss: 7.038543844828382e-05
step: 320, loss: 0.0008093043579719961
step: 330, loss: 0.0006226388504728675
step: 340, loss: 0.005424116272479296
step: 350, loss: 0.009177640080451965
step: 360, loss: 0.00014780895435251296
step: 370, loss: 0.0030788728035986423
step: 380, loss: 0.0053556389175355434
step: 390, loss: 9.968948870664462e-05
step: 400, loss: 0.003738582134246826
step: 410, loss: 0.0016764930915087461
step: 420, loss: 0.00021606290829367936
step: 430, loss: 0.00025567770353518426
step: 440, loss: 0.0215010903775692
step: 450, loss: 0.00013243293506093323
step: 460, loss: 0.00043712768820114434
step: 470, loss: 0.0013302196748554707
step: 480, loss: 0.00022646845900453627
step: 490, loss: 0.00040176272159442306
step: 500, loss: 0.015670614317059517
step: 510, loss: 0.0009441928123123944
step: 520, loss: 0.0001996196515392512
step: 530, loss: 0.0003489409573376179
/usr/local/lib/python3.7/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=0.9459211732355638, f1=0.9402099497946144,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001991492317756638
step: 10, loss: 0.0009702949901111424
step: 20, loss: 0.0003869846696034074
step: 30, loss: 0.0009349590982310474
step: 40, loss: 0.0005864495178684592
step: 50, loss: 0.014530974440276623

```

step: 60, loss: 0.011203448288142681
step: 70, loss: 9.653101733420044e-05
step: 80, loss: 0.0001958527573151514
step: 90, loss: 0.0019336718833073974
step: 100, loss: 0.0017936816439032555
step: 110, loss: 0.00037188531132414937
step: 120, loss: 0.0003167755203321576
step: 130, loss: 0.0002031094627454877
step: 140, loss: 0.002957425080239773
step: 150, loss: 0.00017052942712325603
step: 160, loss: 0.00033934038947336376
step: 170, loss: 0.0003592686844058335
step: 180, loss: 0.00025653079501353204
step: 190, loss: 0.00362038379535079
step: 200, loss: 0.041147224605083466
step: 210, loss: 0.0001784062769729644
step: 220, loss: 0.00265707285143435
step: 230, loss: 0.005182649940252304
step: 240, loss: 0.00015217789041344076
step: 250, loss: 0.00029001987422816455
step: 260, loss: 0.0004880344495177269
step: 270, loss: 0.00589360948652029
step: 280, loss: 8.024153794394806e-05
step: 290, loss: 9.852561197476462e-05
step: 300, loss: 0.00010172634210903198
step: 310, loss: 0.00020364268857520074
step: 320, loss: 0.0005452401819638908
step: 330, loss: 0.0014396774349734187
step: 340, loss: 0.06376703083515167
step: 350, loss: 7.103584357537329e-05
step: 360, loss: 0.0032751422841101885
step: 370, loss: 0.0007950367289595306
step: 380, loss: 5.2692292229039595e-05
step: 390, loss: 0.03982973471283913
step: 400, loss: 0.0009208521223627031
step: 410, loss: 0.002378361765295267
step: 420, loss: 0.0001632860949030146
step: 430, loss: 0.0002715079172048718
step: 440, loss: 0.0004667082102969289
step: 450, loss: 0.00183597591239959
step: 460, loss: 0.009084868244826794
step: 470, loss: 0.0002010049793170765
step: 480, loss: 0.005787254311144352
step: 490, loss: 0.0007268309127539396
step: 500, loss: 0.0027551765087991953
step: 510, loss: 0.000979659496806562
step: 520, loss: 0.00015490036457777023
step: 530, loss: 0.0001591587788425386

```

/usr/local/lib/python3.7/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=0.9503745318352059, f1=0.9441860465116279,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010490157874301076
step: 10, loss: 3.48101784766186e-05
step: 20, loss: 0.07298512756824493
step: 30, loss: 0.001545611536130309
step: 40, loss: 0.011645698919892311
step: 50, loss: 0.046474359929561615
step: 60, loss: 0.001939338748343289
step: 70, loss: 3.849171844194643e-05
step: 80, loss: 0.00023092330957297236
step: 90, loss: 0.05902951955795288
step: 100, loss: 0.06410901248455048
step: 110, loss: 0.02472504787147045
step: 120, loss: 0.002326854970306158
step: 130, loss: 0.0012855351669713855
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.0007789530209265649
step: 150, loss: 0.00010329642100259662
step: 160, loss: 6.669596041319892e-05
step: 170, loss: 0.0008497737580910325
step: 180, loss: 0.0017756057204678655
step: 190, loss: 0.0020342839416116476
step: 200, loss: 0.0011827433481812477
step: 210, loss: 0.00040078122401610017
step: 220, loss: 0.00013625760038848966
step: 230, loss: 0.00031797419069334865
step: 240, loss: 0.00037925090873613954
step: 250, loss: 0.00015336346405092627
step: 260, loss: 7.630597247043625e-05
step: 270, loss: 8.08781769592315e-05
step: 280, loss: 0.0003818203986156732
step: 290, loss: 0.0006793455104343593
step: 300, loss: 0.0003454066172707826
step: 310, loss: 0.0002460102259647101
step: 320, loss: 0.00016931822756305337
step: 330, loss: 0.0003220953221898526
step: 340, loss: 0.00020677538122981787
step: 350, loss: 0.0008303962531499565
step: 360, loss: 0.0006565109943039715
step: 370, loss: 0.000436586095020175
step: 380, loss: 0.00013409979874268174

```

```

step: 390, loss: 0.051793262362480164
step: 400, loss: 0.025708265602588654
step: 410, loss: 0.006362518295645714
step: 420, loss: 0.005110916681587696
step: 430, loss: 0.0033544725738465786
step: 440, loss: 0.003922248259186745
step: 450, loss: 0.005596327595412731
step: 460, loss: 0.00016942991351243109
step: 470, loss: 3.776390803977847e-05
step: 480, loss: 0.0030100068543106318
step: 490, loss: 0.083268903195858
step: 500, loss: 0.0005153638194315135
step: 510, loss: 0.0016244559083133936
step: 520, loss: 0.00022717863612342626
step: 530, loss: 0.00011900727986358106
/usr/local/lib/python3.7/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=0.9483960948396094, f1=0.9416126042632066,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023266011849045753
step: 10, loss: 3.430768265388906e-05
step: 20, loss: 8.144183811964467e-05
step: 30, loss: 0.0005808009300380945
step: 40, loss: 9.277352364733815e-05
step: 50, loss: 0.00043837656266987324
step: 60, loss: 4.770656596519984e-05
step: 70, loss: 0.00017456000205129385
step: 80, loss: 0.00030303894891403615
step: 90, loss: 0.015820475295186043
step: 100, loss: 0.0006934436969459057
step: 110, loss: 0.005243867170065641
step: 120, loss: 0.011934943497180939
step: 130, loss: 6.111004768172279e-05
step: 140, loss: 0.00011347654071869329
step: 150, loss: 0.0068790968507528305
step: 160, loss: 7.518385973526165e-05
step: 170, loss: 0.07760997116565704
step: 180, loss: 0.0007727813208475709
step: 190, loss: 0.00043385126627981663
step: 200, loss: 0.00020922553085256368
step: 210, loss: 0.00010618045052979141
step: 220, loss: 0.0004961060476489365
step: 230, loss: 0.007089188322424889
step: 240, loss: 0.00021400254627224058

```

```

step: 250, loss: 0.0008308527758345008
step: 260, loss: 5.593184687313624e-05
step: 270, loss: 0.00012883497402071953
step: 280, loss: 3.486322384560481e-05
step: 290, loss: 0.0002862693218048662
step: 300, loss: 3.355132503202185e-05
step: 310, loss: 0.0012942035682499409
step: 320, loss: 0.0016853719716891646
step: 330, loss: 0.0006040601874701679
step: 340, loss: 0.00018101916066370904
step: 350, loss: 0.00015763995179440826
step: 360, loss: 1.834300383052323e-05
step: 370, loss: 0.0004584863781929016
step: 380, loss: 0.001110823592171073
step: 390, loss: 0.0024359964299947023
step: 400, loss: 0.006809047888964415
step: 410, loss: 0.0022778783459216356
step: 420, loss: 0.00024593426496721804
step: 430, loss: 0.0009529481758363545
step: 440, loss: 0.00023200473515316844
step: 450, loss: 0.001095209619961679
step: 460, loss: 0.001427658717148006
step: 470, loss: 5.082565621705726e-05
step: 480, loss: 0.0032558089587837458
step: 490, loss: 0.00017461877723690122
step: 500, loss: 0.00015417025133501738
step: 510, loss: 0.0011454926570877433
step: 520, loss: 0.0004524343821685761
step: 530, loss: 5.6282915465999395e-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=0.9489939167056621, f1=0.942271880819367,
best_f1=0.9478584729981377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011278021847829223
step: 10, loss: 4.4409112888388336e-05
step: 20, loss: 4.070721479365602e-05
step: 30, loss: 6.10974311712198e-05
step: 40, loss: 0.0032082207035273314
step: 50, loss: 0.00435671815648675
step: 60, loss: 4.578856533044018e-05
step: 70, loss: 0.0007831549155525863
step: 80, loss: 3.498015212244354e-05
step: 90, loss: 0.0008875406347215176
step: 100, loss: 0.00026546759181655943

```

```
step: 110, loss: 8.86990746948868e-05
step: 120, loss: 0.00011535413068486378
step: 130, loss: 0.009545999579131603
step: 140, loss: 0.0006530214450322092
step: 150, loss: 4.2098825360881165e-05
step: 160, loss: 0.0004432386776898056
step: 170, loss: 0.0004449469852261245
step: 180, loss: 3.730222670128569e-05
step: 190, loss: 0.00011255766003159806
step: 200, loss: 0.00015093307592906058
step: 210, loss: 2.664589192136191e-05
step: 220, loss: 4.9069509259425104e-05
step: 230, loss: 6.512930849567056e-05
step: 240, loss: 3.380842463229783e-05
step: 250, loss: 0.0002390628942521289
step: 260, loss: 0.00040539828478358686
step: 270, loss: 3.6374614865053445e-05
step: 280, loss: 3.597045360947959e-05
step: 290, loss: 0.00028766324976459146
step: 300, loss: 0.00045147346099838614
step: 310, loss: 0.0006470952648669481
step: 320, loss: 0.0007062669610604644
step: 330, loss: 0.0014585148310288787
step: 340, loss: 0.00012168389366706833
step: 350, loss: 4.9427540943725035e-05
step: 360, loss: 0.0015856800600886345
step: 370, loss: 5.0445236411178485e-05
step: 380, loss: 0.0001546367711853236
step: 390, loss: 0.05102742090821266
step: 400, loss: 0.0006198876653797925
step: 410, loss: 0.00021703157108277082
step: 420, loss: 0.0020448516588658094
step: 430, loss: 6.523005868075415e-05
step: 440, loss: 0.0007282737642526627
step: 450, loss: 0.00038205095916055143
step: 460, loss: 0.00020112488709855825
step: 470, loss: 0.00018692275625653565
step: 480, loss: 6.980120087973773e-05
step: 490, loss: 2.4075701730907895e-05
step: 500, loss: 3.652182931546122e-05
step: 510, loss: 0.0005452424520626664
step: 520, loss: 0.0005568405613303185
step: 530, loss: 0.0009446663898415864
/usr/local/lib/python3.7/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=0.9492957746478873, f1=0.9419475655430711,
best_f1=0.9478584729981377
```

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

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

step: 0, loss: 2.9249549697851762e-05
step: 10, loss: 4.646229717764072e-05
step: 20, loss: 7.660298433620483e-05
step: 30, loss: 0.0012958933366462588
step: 40, loss: 0.05144347622990608
step: 50, loss: 0.0020975670777261257
step: 60, loss: 3.211821967852302e-05
step: 70, loss: 9.55371288000606e-05
step: 80, loss: 4.482800795813091e-05
step: 90, loss: 2.9234715839265846e-05
step: 100, loss: 0.0003117061860393733
step: 110, loss: 0.0007000619661994278
step: 120, loss: 0.0007058162009343505
step: 130, loss: 0.08925814181566238
step: 140, loss: 0.09758517146110535
step: 150, loss: 3.6727506085298955e-05
step: 160, loss: 0.0006594073493033648
step: 170, loss: 0.00021515310800168663
step: 180, loss: 4.273282320355065e-05
step: 190, loss: 5.989641067571938e-05
step: 200, loss: 0.00013513177691493183
step: 210, loss: 0.001124502974562347
step: 220, loss: 0.0001351833634544164
step: 230, loss: 0.031005680561065674
step: 240, loss: 0.00037244067061692476
step: 250, loss: 0.00025857670698314905
step: 260, loss: 0.05957895889878273
step: 270, loss: 2.1017562175984494e-05
step: 280, loss: 3.658626519609243e-05
step: 290, loss: 1.82684507308295e-05
step: 300, loss: 2.1140525859664194e-05
step: 310, loss: 0.0005726257222704589
step: 320, loss: 0.0008996442193165421
step: 330, loss: 0.00033710006391629577
step: 340, loss: 9.31454123929143e-05
step: 350, loss: 3.084369018324651e-05
step: 360, loss: 0.0010365432826802135
step: 370, loss: 0.00027910401695407927
step: 380, loss: 0.0001223207073053345
step: 390, loss: 0.0009365449659526348
step: 400, loss: 9.392362699145451e-05
step: 410, loss: 0.002554667182266712
step: 420, loss: 0.0004869991971645504
step: 430, loss: 0.0006375133525580168
step: 440, loss: 0.002392265247181058


```

step: 450, loss: 8.654356497572735e-05
step: 460, loss: 0.00023721106117591262
step: 470, loss: 0.003387260949239135
step: 480, loss: 9.092238906305283e-05
step: 490, loss: 1.5388979591079988e-05
step: 500, loss: 7.804028427926823e-05
step: 510, loss: 0.0001586749858688563
step: 520, loss: 0.00012174038420198485
step: 530, loss: 0.0004625438596121967
/usr/local/lib/python3.7/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=0.9480519480519481, f1=0.9440591770688858,
best_f1=0.9478584729981377
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 [00:24, 235.13it/s]
load_f1 = 0.9481065918653577
real_f1 = 0.9495327102803738
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, 237.39it/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/
```

```
Downloading: 100% 433/433 [00:00<00:00, 378kB/s]
```

```
Downloading: 100% 232k/232k [00:00<00:00, 352kB/s]
```

```
Downloading: 100% 440M/440M [00:06<00:00, 64.6MB/s]
```

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

```
Defaults for this optimization level are:
```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5518725514411926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.42553191489361697, f1=0.38596491228070184,
best_f1=0.38596491228070184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48599323630332947
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5714285714285714, f1=0.4, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39741137623786926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6060606060606061, f1=0.4615384615384615,
best_f1=0.4615384615384615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30297476053237915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7407407407407408, 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: 0.1604718416929245
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7586206896551724, f1=0.7096774193548386,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.12097766250371933
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8, f1=0.7692307692307692, best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012119767256081104
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.7878787878787878,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0040742093697190285
/usr/local/lib/python3.7/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=0.8, f1=0.7878787878787878, best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00165286916308105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7741935483870968, f1=0.7586206896551724,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024720525834709406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.7857142857142857,
best_f1=0.7857142857142857

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016887533711269498
/usr/local/lib/python3.7/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=0.8235294117647058, f1=0.8, best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004320623353123665
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7333333333333334,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014611176447942853
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8275862068965518, f1=0.8148148148148148,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010858398163691163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8275862068965518, f1=0.8148148148148148,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014564072480425239
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8275862068965518, f1=0.8148148148148148,
best_f1=0.7857142857142857
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, 127354.58it/s]
/usr/local/lib/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 = 0.8125000000000001
real_f1 = 0.8125000000000001
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.31it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

```

```

--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

```
step: 0, loss: 0.6302457451820374
step: 10, loss: 0.5886079668998718
step: 20, loss: 0.29834818840026855
step: 30, loss: 0.15624572336673737
step: 40, loss: 0.14196379482746124
step: 50, loss: 0.03251568228006363
step: 60, loss: 0.06347126513719559
step: 70, loss: 0.18352261185646057
step: 80, loss: 0.08455231040716171
```

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

```
step: 90, loss: 0.12522467970848083
step: 100, loss: 0.004396000877022743
step: 110, loss: 0.28962036967277527
step: 120, loss: 0.02514631114900112
step: 130, loss: 0.04731007665395737
step: 140, loss: 0.006429431028664112
step: 150, loss: 0.009417803958058357
step: 160, loss: 0.0038452818989753723
step: 170, loss: 0.031536001712083817
step: 180, loss: 0.002264644019305706
step: 190, loss: 0.0219890084117651
step: 200, loss: 0.02154037542641163
step: 210, loss: 0.005474342964589596
step: 220, loss: 0.0034289562609046698
step: 230, loss: 0.00579547043889761
```

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

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

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

```
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=0.9876819708846584, f1=0.9887133182844244,
best_f1=0.9887133182844244
```

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

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

```
step: 0, loss: 0.002411578083410859
step: 10, loss: 0.0027914198581129313
step: 20, loss: 0.06733465194702148
step: 30, loss: 0.16107147932052612
```

```

step: 40, loss: 0.016194507479667664
step: 50, loss: 0.007571198046207428
step: 60, loss: 0.003706967690959573
step: 70, loss: 0.033857233822345734
step: 80, loss: 0.0016140458174049854
step: 90, loss: 0.0025864967610687017
step: 100, loss: 0.0020958047825843096
step: 110, loss: 0.024697648361325264
step: 120, loss: 0.0017604902386665344
step: 130, loss: 0.08585141599178314
step: 140, loss: 0.002788340440019965
step: 150, loss: 0.02377636916935444
step: 160, loss: 0.006519177462905645
step: 170, loss: 0.0017472266918048263
step: 180, loss: 0.002007711213082075
step: 190, loss: 0.004525036085397005
step: 200, loss: 0.0016661931294947863
step: 210, loss: 0.0009251798619516194
step: 220, loss: 0.019022872671484947
step: 230, loss: 0.0029229559004306793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.990990990990991, f1=0.9852774631936579,
best_f1=0.9852774631936579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017093877540901303
step: 10, loss: 0.0016262826975435019
step: 20, loss: 0.0015352906193584204
step: 30, loss: 0.01951574720442295
step: 40, loss: 0.11299920827150345
step: 50, loss: 0.006149137392640114
step: 60, loss: 0.0013431155821308494
step: 70, loss: 0.0063063958659768105
step: 80, loss: 0.00043947077938355505
step: 90, loss: 0.0017075024079531431
step: 100, loss: 0.0004689568595495075
step: 110, loss: 0.00036212365375831723
step: 120, loss: 0.12885622680187225
step: 130, loss: 0.0008184614707715809
step: 140, loss: 0.004994823597371578
step: 150, loss: 0.002985966159030795

```

```

step: 160, loss: 0.0037629283033311367
step: 170, loss: 0.007592442911118269
step: 180, loss: 0.004353497643023729
step: 190, loss: 0.001400722423568368
step: 200, loss: 0.002606509253382683
step: 210, loss: 0.0008885898860171437
step: 220, loss: 0.001461595413275063
step: 230, loss: 0.00037865390186198056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910112359550561, f1=0.9909706546275394,
best_f1=0.9909706546275394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000772722705733031
step: 10, loss: 0.0002783830859698355
step: 20, loss: 0.00045275368029251695
step: 30, loss: 0.0005213454714976251
step: 40, loss: 0.008829929865896702
step: 50, loss: 0.0011121396673843265
step: 60, loss: 0.000560833839699626
step: 70, loss: 0.0005482195410877466
step: 80, loss: 0.0017007034039124846
step: 90, loss: 0.0009575698641128838
step: 100, loss: 0.00027486542239785194
step: 110, loss: 0.0010401209583505988
step: 120, loss: 0.017367811873555183
step: 130, loss: 0.004004962742328644
step: 140, loss: 0.0014896338107064366
step: 150, loss: 0.07241205126047134
step: 160, loss: 0.002721393946558237
step: 170, loss: 0.005014338064938784
step: 180, loss: 0.0005078729591332376
step: 190, loss: 0.0033594886772334576
step: 200, loss: 0.0004147579602431506
step: 210, loss: 0.15176896750926971
step: 220, loss: 0.00032488195574842393
step: 230, loss: 0.001873224857263267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.992108229988726, f1=0.9875706214689265,
best_f1=0.9875706214689265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001186818815767765
step: 10, loss: 0.0009239138453267515
step: 20, loss: 0.0013170576421543956
step: 30, loss: 0.00020070966274943203
step: 40, loss: 0.0004711913352366537
step: 50, loss: 0.0008037008810788393
step: 60, loss: 0.0006356259691528976
step: 70, loss: 0.0008710872498340905
step: 80, loss: 0.0006151836714707315
step: 90, loss: 0.0006004619062878191
step: 100, loss: 0.012177025899291039
step: 110, loss: 0.0002210458042100072
step: 120, loss: 0.00011445596464909613
step: 130, loss: 0.0006456612609326839
step: 140, loss: 0.0005440996028482914
step: 150, loss: 0.004201121628284454
step: 160, loss: 0.0010827959049493074
step: 170, loss: 0.008052601478993893
step: 180, loss: 0.0020439294166862965
step: 190, loss: 0.001718707848340273
step: 200, loss: 0.00017289808602072299
step: 210, loss: 0.00013053578732069582
step: 220, loss: 0.0003112034755758941
step: 230, loss: 0.00018864356388803571
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.987598647125141,
best_f1=0.9875706214689265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004140599921811372
step: 10, loss: 0.0003640539071056992
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.1799352914094925
step: 30, loss: 0.0002468037127982825
step: 40, loss: 0.00021083377941977233
step: 50, loss: 0.00043754506623372436
step: 60, loss: 0.00010332056262996048
```

```

step: 70, loss: 0.0005240442114882171
step: 80, loss: 0.002213912084698677
step: 90, loss: 0.002219748916104436
step: 100, loss: 0.06712185591459274
step: 110, loss: 0.0003500574384815991
step: 120, loss: 0.004656959790736437
step: 130, loss: 0.00034884450724348426
step: 140, loss: 0.002071941737085581
step: 150, loss: 0.010508962906897068
step: 160, loss: 0.003119958797469735
step: 170, loss: 0.00023145087470766157
step: 180, loss: 0.06697990000247955
step: 190, loss: 0.034768715500831604
step: 200, loss: 0.00017912806652020663
step: 210, loss: 0.00017977232346311212
step: 220, loss: 0.00017473682237323374
step: 230, loss: 0.04864272475242615
/usr/local/lib/python3.7/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=0.990990990990991, f1=0.9865168539325843,
best_f1=0.9875706214689265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028782677836716175
step: 10, loss: 9.110131213674322e-05
step: 20, loss: 0.0033147132489830256
step: 30, loss: 0.0013038848992437124
step: 40, loss: 0.025440705940127373
step: 50, loss: 0.0001605315919732675
step: 60, loss: 0.002448085229843855
step: 70, loss: 0.020610613748431206
step: 80, loss: 0.0005000985693186522
step: 90, loss: 6.503076292574406e-05
step: 100, loss: 9.475665137870237e-05
step: 110, loss: 0.0001977833017008379
step: 120, loss: 0.00013801601016893983
step: 130, loss: 9.304812556365505e-05
step: 140, loss: 0.00010578351793810725
step: 150, loss: 0.0016820139717310667
step: 160, loss: 0.08171665668487549
step: 170, loss: 0.0019368738867342472
step: 180, loss: 0.003650303464382887
step: 190, loss: 0.00033762483508326113
step: 200, loss: 0.06922218203544617
step: 210, loss: 7.946640835143626e-05
step: 220, loss: 0.00019417090516071767

```

```

step: 230, loss: 0.0006057795253582299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921436588103255, f1=0.978865406006674,
best_f1=0.978865406006674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016731540381442755
step: 10, loss: 0.002154156332835555
step: 20, loss: 0.0007331817178055644
step: 30, loss: 0.00020632566884160042
step: 40, loss: 0.000305812805891037
step: 50, loss: 0.00033938157139346004
step: 60, loss: 7.743716560071334e-05
step: 70, loss: 0.00014017827925272286
step: 80, loss: 0.000175897337612696
step: 90, loss: 9.565698564983904e-05
step: 100, loss: 0.00030786238494329154
step: 110, loss: 0.09405473619699478
step: 120, loss: 0.00017844533431343734
step: 130, loss: 0.00037201461964286864
step: 140, loss: 6.851273064967245e-05
step: 150, loss: 0.0002673301787581295
step: 160, loss: 0.00023196243273559958
step: 170, loss: 6.943225162103772e-05
step: 180, loss: 9.78664611466229e-05
step: 190, loss: 0.001627582823857665
step: 200, loss: 6.059229781385511e-05
step: 210, loss: 0.00011557328252820298
step: 220, loss: 0.0006944357301108539
step: 230, loss: 0.0004207183956168592
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.9811320754716982,
best_f1=0.978865406006674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.3572636059252545e-05
step: 10, loss: 0.000581117463298142
step: 20, loss: 0.0004640293773263693

```

```

step: 30, loss: 6.581744673894718e-05
step: 40, loss: 0.023688357323408127
step: 50, loss: 5.253743438515812e-05
step: 60, loss: 0.0001875428279163316
step: 70, loss: 8.81223168107681e-05
step: 80, loss: 6.487136124633253e-05
step: 90, loss: 7.454185106325895e-05
step: 100, loss: 0.00011786993127316236
step: 110, loss: 7.21719698049128e-05
step: 120, loss: 4.1900682845152915e-05
step: 130, loss: 6.431736255763099e-05
step: 140, loss: 3.640607974375598e-05
step: 150, loss: 0.00010738804849097505
step: 160, loss: 5.6768101785564795e-05
step: 170, loss: 0.00010534885950619355
step: 180, loss: 0.00038182755815796554
step: 190, loss: 0.00027097552083432674
step: 200, loss: 0.00010054872109321877
step: 210, loss: 9.956552821677178e-05
step: 220, loss: 4.743957833852619e-05
step: 230, loss: 0.000106331157439854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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=0.9932432432432432, f1=0.9887640449438202,
best_f1=0.9887640449438202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010959932842524722
step: 10, loss: 5.787761620013043e-05
step: 20, loss: 0.00013465066149365157
step: 30, loss: 7.083761011017486e-05
step: 40, loss: 5.689080717274919e-05
step: 50, loss: 4.6481181925628334e-05
step: 60, loss: 0.0033375907223671675
step: 70, loss: 0.00014435221964959055
step: 80, loss: 4.961082959198393e-05
step: 90, loss: 0.0001356756838504225
step: 100, loss: 4.071238072356209e-05
step: 110, loss: 0.00010035577724920586
step: 120, loss: 0.000293406134005636
step: 130, loss: 9.671373845776543e-05
step: 140, loss: 0.03577980399131775

```

```

step: 150, loss: 0.016378087922930717
step: 160, loss: 3.041999116248917e-05
step: 170, loss: 0.0002566082403063774
step: 180, loss: 0.00011173867096658796
step: 190, loss: 0.023093730211257935
step: 200, loss: 0.0001917546906042844
step: 210, loss: 4.2199109884677455e-05
step: 220, loss: 3.636151450336911e-05
step: 230, loss: 0.0001809947716537863
/usr/local/lib/python3.7/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=0.9932432432432432, f1=0.9876543209876544,
best_f1=0.9887640449438202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.8212627259781584e-05
step: 10, loss: 5.263639832264744e-05
step: 20, loss: 0.00010230633051833138
step: 30, loss: 0.005330681800842285
step: 40, loss: 0.0003772190830204636
step: 50, loss: 3.695722625707276e-05
step: 60, loss: 5.301445708028041e-05
step: 70, loss: 6.458810094045475e-05
step: 80, loss: 2.7860874979523942e-05
step: 90, loss: 6.26705150352791e-05
step: 100, loss: 4.4440046622185037e-05
step: 110, loss: 0.024246202781796455
step: 120, loss: 3.17756857839413e-05
step: 130, loss: 2.3949412934598513e-05
step: 140, loss: 5.919485192862339e-05
step: 150, loss: 0.041067223995923996
step: 160, loss: 3.383202783879824e-05
step: 170, loss: 0.01743144914507866
step: 180, loss: 3.559013202902861e-05
step: 190, loss: 3.16044497594703e-05
step: 200, loss: 0.0001455270394217223
step: 210, loss: 3.14590688503813e-05
step: 220, loss: 3.1392250093631446e-05
step: 230, loss: 3.781715713557787e-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=0.9920903954802259, f1=0.9898534385569334,
best_f1=0.9887640449438202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.817879279376939e-05
step: 10, loss: 3.436462429817766e-05
step: 20, loss: 0.00011421171802794561
step: 30, loss: 3.378012843313627e-05
step: 40, loss: 5.8279478253098205e-05
step: 50, loss: 9.353589848615229e-05
step: 60, loss: 0.0005392354214563966
step: 70, loss: 3.444984395173378e-05
step: 80, loss: 6.585709343198687e-05
step: 90, loss: 2.5070708943530917e-05
step: 100, loss: 4.270023782737553e-05
step: 110, loss: 3.116480729659088e-05
step: 120, loss: 4.269354394637048e-05
step: 130, loss: 3.771358387893997e-05
step: 140, loss: 0.0031360217835754156
step: 150, loss: 0.0005453184712678194
step: 160, loss: 3.221174847567454e-05
step: 170, loss: 4.094181349501014e-05
step: 180, loss: 0.025313114747405052
step: 190, loss: 0.0002239874447695911
step: 200, loss: 2.1945232219877653e-05
step: 210, loss: 2.7581427275436e-05
step: 220, loss: 2.8046977604390122e-05
step: 230, loss: 0.055112890899181366
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9943883277216611, f1=0.9876819708846584,
best_f1=0.9876819708846584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.959051173296757e-05
step: 10, loss: 0.00010529175051487982
step: 20, loss: 5.885290738660842e-05
step: 30, loss: 4.028375769848935e-05
step: 40, loss: 3.370525155332871e-05
step: 50, loss: 5.638349102810025e-05
step: 60, loss: 3.293414556537755e-05
step: 70, loss: 2.3617831175215542e-05
step: 80, loss: 2.771919571387116e-05
step: 90, loss: 5.936520756222308e-05
step: 100, loss: 2.1155587091925554e-05

```

```

step: 110, loss: 0.0147119564935565
step: 120, loss: 0.03388224169611931
step: 130, loss: 3.318770177429542e-05
step: 140, loss: 3.0229834010242485e-05
step: 150, loss: 3.095261490670964e-05
step: 160, loss: 4.9280508392257616e-05
step: 170, loss: 0.0001302607706747949
step: 180, loss: 3.709020165842958e-05
step: 190, loss: 2.4124405172187835e-05
step: 200, loss: 4.2769032006617635e-05
step: 210, loss: 2.016094367718324e-05
step: 220, loss: 2.6299960154574364e-05
step: 230, loss: 7.854281284380704e-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=0.9943883277216611, f1=0.9865470852017937,
best_f1=0.9876819708846584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7242327632848173e-05
step: 10, loss: 0.0001614254288142547
step: 20, loss: 2.5633180484874174e-05
step: 30, loss: 0.00011321877536829561
step: 40, loss: 0.016155559569597244
step: 50, loss: 2.8713902793242596e-05
step: 60, loss: 2.8378521164995618e-05
step: 70, loss: 3.6613375414162874e-05
step: 80, loss: 4.115311458008364e-05
step: 90, loss: 3.850922439596616e-05
step: 100, loss: 2.468327511451207e-05
step: 110, loss: 0.0001401982590323314
step: 120, loss: 1.69200557138538e-05
step: 130, loss: 2.4970137019408867e-05
step: 140, loss: 4.6985354856587946e-05
step: 150, loss: 2.370356560277287e-05
step: 160, loss: 0.002233977196738124
step: 170, loss: 0.00015750255261082202
step: 180, loss: 0.0001215977463289164
step: 190, loss: 5.101895294501446e-05
step: 200, loss: 2.3364553271676414e-05
step: 210, loss: 3.491944880806841e-05
step: 220, loss: 4.6984958316897973e-05
step: 230, loss: 4.125719715375453e-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=0.9943883277216611, f1=0.9876819708846584,
best_f1=0.9876819708846584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.0015655536553822e-05
step: 10, loss: 1.8614982764120214e-05
step: 20, loss: 2.653098817972932e-05
step: 30, loss: 3.6598154110834e-05
step: 40, loss: 6.487399514298886e-05
step: 50, loss: 0.031183674931526184
step: 60, loss: 3.5906239645555615e-05
step: 70, loss: 7.45165380067192e-05
step: 80, loss: 3.860550350509584e-05
step: 90, loss: 5.007370054954663e-05
step: 100, loss: 4.125027044210583e-05
step: 110, loss: 2.398656761215534e-05
step: 120, loss: 6.362768908729777e-05
step: 130, loss: 0.022305116057395935
step: 140, loss: 2.1878258849028498e-05
step: 150, loss: 0.00012326179421506822
step: 160, loss: 2.1811018086737022e-05
step: 170, loss: 2.0194505850668065e-05
step: 180, loss: 6.179218325996771e-05
step: 190, loss: 0.0002680308243725449
step: 200, loss: 3.509823727654293e-05
step: 210, loss: 2.653095725690946e-05
step: 220, loss: 5.567304106079973e-05
step: 230, loss: 2.3558224711450748e-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=0.9943883277216611, f1=0.9865771812080537,
best_f1=0.9876819708846584
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, 161.90it/s]
load_f1 = 0.9943883277216611
real_f1 = 0.9932584269662922
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, 179.83it/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: 0.6244491338729858
```

```
step: 10, loss: 0.529644250869751
```

step: 20, loss: 0.4983280897140503
step: 30, loss: 0.08074798434972763
step: 40, loss: 0.17249082028865814
step: 50, loss: 0.17855846881866455
step: 60, loss: 0.15022696554660797
step: 70, loss: 0.11631770431995392
step: 80, loss: 0.03404606506228447
step: 90, loss: 0.03786586970090866
step: 100, loss: 0.005029382649809122
step: 110, loss: 0.1430402398109436
step: 120, loss: 0.07953217625617981
step: 130, loss: 0.1283522993326187
step: 140, loss: 0.1444178968667984
step: 150, loss: 0.01470216829329729
step: 160, loss: 0.016260143369436264
step: 170, loss: 0.14958958327770233
step: 180, loss: 0.08089753240346909
step: 190, loss: 0.006884570233523846
step: 200, loss: 0.15607218444347382
step: 210, loss: 0.15640299022197723
step: 220, loss: 0.27127164602279663
step: 230, loss: 0.1498737931251526
step: 240, loss: 0.0424480177462101
step: 250, loss: 0.0076844762079417706
step: 260, loss: 0.06217306852340698
step: 270, loss: 0.005748485680669546
step: 280, loss: 0.02668857015669346
step: 290, loss: 0.10129795968532562
step: 300, loss: 0.039125312119722366
step: 310, loss: 0.1628459244966507
step: 320, loss: 0.1323491334915161
step: 330, loss: 0.004794859327375889
step: 340, loss: 0.01018694881349802
step: 350, loss: 0.022047018632292747
step: 360, loss: 0.03258443996310234
step: 370, loss: 0.03380132094025612
step: 380, loss: 0.014840507879853249
step: 390, loss: 0.1534995585680008
step: 400, loss: 0.18696671724319458
step: 410, loss: 0.039084360003471375
step: 420, loss: 0.023820916190743446
step: 430, loss: 0.18768543004989624
step: 440, loss: 0.013023595325648785
step: 450, loss: 0.004042275249958038
step: 460, loss: 0.0049414439126849174
step: 470, loss: 0.18753905594348907
step: 480, loss: 0.027054302394390106
step: 490, loss: 0.05983813852071762

```

step: 500, loss: 0.06271877139806747
step: 510, loss: 0.04088513180613518
step: 520, loss: 0.03144027292728424
step: 530, loss: 0.002750096609815955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.949343339587242, f1=0.9406264609630668,
best_f1=0.9406264609630668
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14803087711334229
step: 10, loss: 0.025125710293650627
step: 20, loss: 0.044326718896627426
step: 30, loss: 0.02367766574025154
step: 40, loss: 0.1210230141878128
step: 50, loss: 0.1018688753247261
step: 60, loss: 0.0032741804607212543
step: 70, loss: 0.009260416962206364
step: 80, loss: 0.03194589540362358
step: 90, loss: 0.020583372563123703
step: 100, loss: 0.009337651543319225
step: 110, loss: 0.005652690306305885
step: 120, loss: 0.06128925085067749
step: 130, loss: 0.029664501547813416
step: 140, loss: 0.10857207328081131
step: 150, loss: 0.055081289261579514
step: 160, loss: 0.012881100177764893
step: 170, loss: 0.018791204318404198
step: 180, loss: 0.07853817194700241
step: 190, loss: 0.07055528461933136
step: 200, loss: 0.011216970160603523
step: 210, loss: 0.056688494980335236
step: 220, loss: 0.09159502387046814
step: 230, loss: 0.01314077153801918
step: 240, loss: 0.019365346059203148
step: 250, loss: 0.03616548329591751
step: 260, loss: 0.0038576710503548384
step: 270, loss: 0.24994434416294098
step: 280, loss: 0.017594676464796066
step: 290, loss: 0.026575002819299698
step: 300, loss: 0.19421252608299255
step: 310, loss: 0.007276302203536034

```

```

step: 320, loss: 0.06677106022834778
step: 330, loss: 0.029133517295122147
step: 340, loss: 0.010354815050959587
step: 350, loss: 0.000738097878638655
step: 360, loss: 0.02036241441965103
step: 370, loss: 0.0744808241724968
step: 380, loss: 0.06301085650920868
step: 390, loss: 0.030799448490142822
step: 400, loss: 0.010570326820015907
step: 410, loss: 0.012173155322670937
step: 420, loss: 0.06265776604413986
step: 430, loss: 0.016973400488495827
step: 440, loss: 0.2179974764585495
step: 450, loss: 0.008886427618563175
step: 460, loss: 0.019510503858327866
step: 470, loss: 0.10083989799022675
step: 480, loss: 0.2719542980194092
step: 490, loss: 0.013583670370280743
step: 500, loss: 0.11531000584363937
step: 510, loss: 0.006830202881246805
step: 520, loss: 0.036819688975811005
step: 530, loss: 0.009214946068823338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9533487297921477, f1=0.9480459770114943,
best_f1=0.9480459770114943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01754174567759037
step: 10, loss: 0.05542021989822388
step: 20, loss: 0.04200151562690735
step: 30, loss: 0.14769478142261505
step: 40, loss: 0.017195187509059906
step: 50, loss: 0.05333670228719711
step: 60, loss: 0.006779715418815613
step: 70, loss: 0.005097134504467249
step: 80, loss: 0.0009871094953268766
step: 90, loss: 0.005864162929356098
step: 100, loss: 0.09438209980726242
step: 110, loss: 0.0032408004626631737
step: 120, loss: 0.001969496253877878
step: 130, loss: 0.0008388530695810914

```



```

step: 140, loss: 0.08609769493341446
step: 150, loss: 0.04779093712568283
step: 160, loss: 0.01632610149681568
step: 170, loss: 0.03933116793632507
step: 180, loss: 0.013823885470628738
step: 190, loss: 0.0010069942800328135
step: 200, loss: 0.003638865891844034
step: 210, loss: 0.0792284905910492
step: 220, loss: 0.06187848746776581
step: 230, loss: 0.04271518066525459
step: 240, loss: 0.0015212380094453692
step: 250, loss: 0.06192995235323906
step: 260, loss: 0.014614799991250038
step: 270, loss: 0.0020713151898235083
step: 280, loss: 0.16593247652053833
step: 290, loss: 0.0015065909828990698
step: 300, loss: 0.00997241772711277
step: 310, loss: 0.009797734208405018
step: 320, loss: 0.031402453780174255
step: 330, loss: 0.010209817439317703
step: 340, loss: 0.013981287367641926
step: 350, loss: 0.005922404583543539
step: 360, loss: 0.029370715841650963
step: 370, loss: 0.0050069186836481094
step: 380, loss: 0.06708141416311264
step: 390, loss: 0.037957869470119476
step: 400, loss: 0.0074227130971848965
step: 410, loss: 0.004649012349545956
step: 420, loss: 0.2500569820404053
step: 430, loss: 0.002569387899711728
step: 440, loss: 0.10908207297325134
step: 450, loss: 0.10224562138319016
step: 460, loss: 0.03852969780564308
step: 470, loss: 0.016243962571024895
step: 480, loss: 0.0020815273746848106
step: 490, loss: 0.012742817401885986
step: 500, loss: 0.005727565847337246
step: 510, loss: 0.0018715558107942343
step: 520, loss: 0.012063520960509777
step: 530, loss: 0.009096705354750156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9569245020842984, f1=0.9458583988894032,

```

```
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017836138606071472
step: 10, loss: 0.002871593227609992
step: 20, loss: 0.0008009398588910699
step: 30, loss: 0.00031586369732394814
step: 40, loss: 0.0016203968552872539
step: 50, loss: 0.0006117871962487698
step: 60, loss: 0.0007486878312192857
step: 70, loss: 0.00036225232179276645
step: 80, loss: 0.017740556970238686
step: 90, loss: 0.02962004952132702
step: 100, loss: 0.0015095864655449986
step: 110, loss: 0.010028442367911339
step: 120, loss: 0.004617678001523018
step: 130, loss: 0.001420289627276361
step: 140, loss: 0.0003480495070107281
step: 150, loss: 0.004744208883494139
step: 160, loss: 0.0030096028931438923
step: 170, loss: 0.001765220076777041
step: 180, loss: 0.0013140999944880605
step: 190, loss: 0.004813011270016432
step: 200, loss: 0.0024135708808898926
step: 210, loss: 0.04008975625038147
step: 220, loss: 0.12145377695560455
step: 230, loss: 0.08123473078012466
step: 240, loss: 0.0022923359647393227
step: 250, loss: 0.04456248879432678
step: 260, loss: 0.0018182331696152687
step: 270, loss: 0.00821913592517376
step: 280, loss: 0.011161201633512974
step: 290, loss: 0.0027065472677350044
step: 300, loss: 0.00018634580192156136
step: 310, loss: 0.011066907085478306
step: 320, loss: 0.031871091574430466
step: 330, loss: 0.03369302675127983
step: 340, loss: 0.0005887735751457512
step: 350, loss: 0.0009313499322161078
step: 360, loss: 0.002677099546417594
step: 370, loss: 0.003437300445511937
step: 380, loss: 0.01300010085105896
step: 390, loss: 0.018963364884257317
step: 400, loss: 0.00035348799428902566
step: 410, loss: 0.0015766811557114124
step: 420, loss: 0.0036517574917525053
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
```

```
step: 430, loss: 0.032669346779584885
step: 440, loss: 0.004187518265098333
step: 450, loss: 0.00751062948256731
step: 460, loss: 0.000365057640010491
step: 470, loss: 0.0035220636054873466
step: 480, loss: 0.02483367919921875
step: 490, loss: 0.0015925879124552011
step: 500, loss: 0.0012037134729325771
step: 510, loss: 0.012665608897805214
step: 520, loss: 0.12981347739696503
step: 530, loss: 0.007472009398043156
/usr/local/lib/python3.7/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=0.9445983379501385, f1=0.9417744916820702,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012829558923840523
step: 10, loss: 0.03158649802207947
step: 20, loss: 0.0036228590179234743
step: 30, loss: 0.06659956276416779
step: 40, loss: 0.0008661944302730262
step: 50, loss: 0.023079315200448036
step: 60, loss: 0.03415418788790703
step: 70, loss: 0.001141160144470632
step: 80, loss: 0.0028424696065485477
step: 90, loss: 0.0006017438136041164
step: 100, loss: 0.008823291398584843
step: 110, loss: 0.0003928770893253386
step: 120, loss: 0.0005143533926457167
step: 130, loss: 0.00024490847135894
step: 140, loss: 0.0009290591697208583
step: 150, loss: 0.0005018568481318653
step: 160, loss: 0.00024554249830543995
step: 170, loss: 0.01463407464325428
step: 180, loss: 0.00021684101375285536
step: 190, loss: 0.00017460093658883125
step: 200, loss: 0.00010819501767400652
step: 210, loss: 0.1564776748418808
step: 220, loss: 0.002253063954412937
step: 230, loss: 0.0026885501574724913
step: 240, loss: 0.0004132137692067772
step: 250, loss: 0.00012348193558864295
step: 260, loss: 0.0006288379663601518
step: 270, loss: 0.00047625359729863703
step: 280, loss: 0.001493309740908444
```

```

step: 290, loss: 0.10089163482189178
step: 300, loss: 0.003165437141433358
step: 310, loss: 0.001573388115502894
step: 320, loss: 0.05611381679773331
step: 330, loss: 0.007591617293655872
step: 340, loss: 0.004500912968069315
step: 350, loss: 0.056646011769771576
step: 360, loss: 0.005386881995946169
step: 370, loss: 0.017369559034705162
step: 380, loss: 0.00022482099302578717
step: 390, loss: 0.0002506282180547714
step: 400, loss: 0.0031494153663516045
step: 410, loss: 0.0006460251752287149
step: 420, loss: 0.008025705814361572
step: 430, loss: 0.006791192106902599
step: 440, loss: 0.023815324530005455
step: 450, loss: 0.0007889072876423597
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 460, loss: 0.15498515963554382
step: 470, loss: 0.00028492778073996305
step: 480, loss: 0.0010037600295618176
step: 490, loss: 0.007748086471110582
step: 500, loss: 0.005273136775940657
step: 510, loss: 0.002582712797448039
step: 520, loss: 0.00027540980954654515
step: 530, loss: 0.019181126728653908
/usr/local/lib/python3.7/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=0.9467680608365018, f1=0.934792955735364,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004191347397863865
step: 10, loss: 9.514619159745052e-05
step: 20, loss: 0.08130878955125809
step: 30, loss: 0.0005137915723025799
step: 40, loss: 0.025532163679599762
step: 50, loss: 0.04318808764219284
step: 60, loss: 0.0007218188839033246
step: 70, loss: 0.0008813437889330089
step: 80, loss: 0.0030841610860079527
step: 90, loss: 0.00017320318147540092
step: 100, loss: 0.0011441364185884595
step: 110, loss: 0.0005801777588203549
step: 120, loss: 0.0014961975393816829
step: 130, loss: 0.0014561852440238

```

```

step: 140, loss: 0.03680827096104622
step: 150, loss: 0.004849869292229414
step: 160, loss: 0.0019603099208325148
step: 170, loss: 0.000961586250923574
step: 180, loss: 0.00564462598413229
step: 190, loss: 0.0007845432846806943
step: 200, loss: 0.0029928251169621944
step: 210, loss: 0.0018692626617848873
step: 220, loss: 0.003251966554671526
step: 230, loss: 0.06321585178375244
step: 240, loss: 0.05473090335726738
step: 250, loss: 0.0005466557922773063
step: 260, loss: 0.00015689381689298898
step: 270, loss: 0.06968111544847488
step: 280, loss: 0.001074864063411951
step: 290, loss: 0.006139862816780806
step: 300, loss: 0.0006696040509268641
step: 310, loss: 0.0005585760227404535
step: 320, loss: 0.00010263500007567927
step: 330, loss: 0.00036220563924871385
step: 340, loss: 0.15329863131046295
step: 350, loss: 0.0023020668886601925
step: 360, loss: 0.013180217705667019
step: 370, loss: 0.0012159976176917553
step: 380, loss: 0.00022921663185115904
step: 390, loss: 0.011953363195061684
step: 400, loss: 0.00022131511650513858
step: 410, loss: 0.00020399526692926884
step: 420, loss: 0.0002594216784927994
step: 430, loss: 0.0007752933306619525
step: 440, loss: 6.116952863521874e-05
step: 450, loss: 0.00014892061881255358
step: 460, loss: 5.8274261391488835e-05
step: 470, loss: 0.00021515344269573689
step: 480, loss: 0.0011735763400793076
step: 490, loss: 0.011564472690224648
step: 500, loss: 0.0023067882284522057
step: 510, loss: 0.0035378076136112213
step: 520, loss: 0.010320623405277729
step: 530, loss: 0.005681220907717943
/usr/local/lib/python3.7/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=0.9535315985130112, f1=0.9473193473193473,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.002639446407556534
step: 10, loss: 0.08609851449728012
step: 20, loss: 0.006193880457431078
step: 30, loss: 0.0006582638015970588
step: 40, loss: 0.00012183303624624386
step: 50, loss: 0.0033709965646266937
step: 60, loss: 0.000522338435985148
step: 70, loss: 0.0003014246467500925
step: 80, loss: 0.03414096683263779
step: 90, loss: 0.055988457053899765
step: 100, loss: 0.02750765159726143
step: 110, loss: 0.004354129079729319
step: 120, loss: 0.0019544842652976513
step: 130, loss: 0.0010546521516516805
step: 140, loss: 0.00033069762866944075
step: 150, loss: 0.0008569885394535959
step: 160, loss: 0.0001268258347408846
step: 170, loss: 0.00041964807314798236
step: 180, loss: 0.006726684048771858
step: 190, loss: 0.0019006957300007343
step: 200, loss: 0.00023113666975405067
step: 210, loss: 0.0003283843689132482
step: 220, loss: 6.334059435175732e-05
step: 230, loss: 0.0025086109526455402
step: 240, loss: 0.00023224795586429536
step: 250, loss: 0.0002914889482781291
step: 260, loss: 9.032394882524386e-05
step: 270, loss: 7.957460184115916e-05
step: 280, loss: 0.00013575449702329934
step: 290, loss: 6.055075573385693e-05
step: 300, loss: 0.0020217217970639467
step: 310, loss: 4.6061464672675356e-05
step: 320, loss: 0.0026403965894132853
step: 330, loss: 0.00034760255948640406
step: 340, loss: 0.015800857916474342
step: 350, loss: 0.005271545145660639
step: 360, loss: 0.0007558173965662718
step: 370, loss: 6.193975423229858e-05
step: 380, loss: 0.010086029767990112
step: 390, loss: 0.00012375880032777786
step: 400, loss: 0.030143359676003456
step: 410, loss: 0.001601995900273323
step: 420, loss: 0.00017392574227415025
step: 430, loss: 4.703291051555425e-05
step: 440, loss: 6.732057954650372e-05
step: 450, loss: 0.0008093537180684507
step: 460, loss: 0.001148201059550047
step: 470, loss: 0.002858813852071762

```

step: 480, loss: 0.0794072300195694
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 490, loss: 0.00011920285760425031
step: 500, loss: 8.232535765273497e-05
step: 510, loss: 0.0006836918764747679
step: 520, loss: 0.019934630021452904
step: 530, loss: 0.00042376748751848936
/usr/local/lib/python3.7/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=0.9444444444444445, f1=0.9432723863103609,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.212078005773947e-05
step: 10, loss: 0.07802638411521912
step: 20, loss: 4.922512380289845e-05
step: 30, loss: 7.57377638365142e-05
step: 40, loss: 0.00011320942576276138
step: 50, loss: 0.038512274622917175
step: 60, loss: 0.0034161871299147606
step: 70, loss: 0.0013311607763171196
step: 80, loss: 0.013700559735298157
step: 90, loss: 0.0018783885752782226
step: 100, loss: 0.0010527889244258404
step: 110, loss: 0.00026953310589306056
step: 120, loss: 0.0014462786493822932
step: 130, loss: 0.0003792117058765143
step: 140, loss: 0.0012796513037756085
step: 150, loss: 0.00013899648911319673
step: 160, loss: 0.015543235465884209
step: 170, loss: 0.00729482714086771
step: 180, loss: 3.58657562173903e-05
step: 190, loss: 0.0015980905154719949
step: 200, loss: 0.001440370106138289
step: 210, loss: 5.848443106515333e-05
step: 220, loss: 0.001774786738678813
step: 230, loss: 0.007922614924609661
step: 240, loss: 5.4197200370253995e-05
step: 250, loss: 0.00023384827363770455
step: 260, loss: 0.00015840516425669193
step: 270, loss: 0.005292095709592104
step: 280, loss: 0.09226406365633011
step: 290, loss: 0.00015060760779306293
step: 300, loss: 0.001132480800151825
step: 310, loss: 0.02440435253083706
step: 320, loss: 0.04724506288766861

```

```

step: 330, loss: 0.004686238244175911
step: 340, loss: 0.0009083659970201552
step: 350, loss: 0.00015257959603331983
step: 360, loss: 0.0017195140244439244
step: 370, loss: 0.0003446867340244353
step: 380, loss: 0.04020242020487785
step: 390, loss: 0.11633933335542679
step: 400, loss: 0.003380089532583952
step: 410, loss: 0.003093732986599207
step: 420, loss: 0.004104550927877426
step: 430, loss: 0.04607045277953148
step: 440, loss: 0.0008502732962369919
step: 450, loss: 0.00020272471010684967
step: 460, loss: 0.00045465247239917517
step: 470, loss: 0.0001767134090187028
step: 480, loss: 0.0006442128214985132
step: 490, loss: 0.041606541723012924
step: 500, loss: 0.000291857635602355
step: 510, loss: 0.00012095545389456674
step: 520, loss: 0.0023025386035442352
step: 530, loss: 0.00031191573361866176
/usr/local/lib/python3.7/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=0.9446254071661238, f1=0.9497464269248502,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013028350658714771
step: 10, loss: 0.00038671656511723995
step: 20, loss: 0.00372281763702631
step: 30, loss: 0.003541052108630538
step: 40, loss: 0.0009309271117672324
step: 50, loss: 0.00013172051694709808
step: 60, loss: 7.845288200769573e-05
step: 70, loss: 0.0012237776536494493
step: 80, loss: 0.0023275804705917835
step: 90, loss: 8.173424430424348e-05
step: 100, loss: 3.453923272900283e-05
step: 110, loss: 7.298224227270111e-05
step: 120, loss: 0.00010541946539888158
step: 130, loss: 0.0002369092107983306
step: 140, loss: 8.328885451192036e-05
step: 150, loss: 0.00035430144635029137
step: 160, loss: 0.0004648083704523742
step: 170, loss: 0.007280570454895496
step: 180, loss: 5.6644039432285354e-05

```



```

step: 190, loss: 0.00011247173824813217
step: 200, loss: 0.001106632873415947
step: 210, loss: 0.00011499031825223938
step: 220, loss: 7.03822443028912e-05
step: 230, loss: 3.615605237428099e-05
step: 240, loss: 0.00016899128968361765
step: 250, loss: 0.0003913216642104089
step: 260, loss: 0.07686332613229752
step: 270, loss: 2.5372268282808363e-05
step: 280, loss: 0.027224067598581314
step: 290, loss: 0.03801339119672775
step: 300, loss: 4.039123814436607e-05
step: 310, loss: 0.047914184629917145
step: 320, loss: 0.008243424817919731
step: 330, loss: 4.844135764869861e-05
step: 340, loss: 6.0368583945091814e-05
step: 350, loss: 0.00021426148305181414
step: 360, loss: 2.5364952307427302e-05
step: 370, loss: 0.00669232290238142
step: 380, loss: 3.2151987397810444e-05
step: 390, loss: 3.10378527501598e-05
step: 400, loss: 4.132973117521033e-05
step: 410, loss: 3.1599789508618414e-05
step: 420, loss: 3.126421506749466e-05
step: 430, loss: 2.417287578282412e-05
step: 440, loss: 0.0018828491447493434
step: 450, loss: 3.8435937312897295e-05
step: 460, loss: 0.00012290528684388846
step: 470, loss: 2.4627232050988823e-05
step: 480, loss: 6.0693666455335915e-05
step: 490, loss: 0.004594534169882536
step: 500, loss: 0.0024967328645288944
step: 510, loss: 0.0007067013066262007
step: 520, loss: 3.953119812649675e-05
step: 530, loss: 0.00014455609198193997
/usr/local/lib/python3.7/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=0.9516279069767443, f1=0.9447004608294931,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.591896089958027e-05
step: 10, loss: 0.00017416865739505738
step: 20, loss: 3.1834995752433315e-05
step: 30, loss: 4.066903056809679e-05
step: 40, loss: 0.0001581040269229561

```

step: 50, loss: 0.0318751186132431
step: 60, loss: 5.7931629271479324e-05
step: 70, loss: 0.00032583417487330735
step: 80, loss: 3.2040279620559886e-05
step: 90, loss: 3.44052241416648e-05
step: 100, loss: 0.00011088964674854651
step: 110, loss: 2.271614175697323e-05
step: 120, loss: 0.0001079954017768614
step: 130, loss: 0.00011336906754877418
step: 140, loss: 0.0029094789642840624
step: 150, loss: 0.0007936538895592093
step: 160, loss: 0.04826021566987038
step: 170, loss: 0.00014469990856014192
step: 180, loss: 0.0022836565040051937
step: 190, loss: 6.082194158807397e-05
step: 200, loss: 0.0002314582816325128
step: 210, loss: 0.00017768441466614604
step: 220, loss: 0.01121798437088728
step: 230, loss: 0.002640328137204051
step: 240, loss: 0.00013816369755659252
step: 250, loss: 5.79239749640692e-05
step: 260, loss: 0.001785462023690343
step: 270, loss: 0.00039241387275978923
step: 280, loss: 0.00023882027016952634
step: 290, loss: 0.00010213301720796153
step: 300, loss: 2.3841121219447814e-05
step: 310, loss: 4.779708251589909e-05
step: 320, loss: 0.0017455036286264658
step: 330, loss: 3.356615707161836e-05
step: 340, loss: 0.013046528212726116
step: 350, loss: 2.7897816835320555e-05
step: 360, loss: 4.8862835683394223e-05
step: 370, loss: 0.0012664921814575791
step: 380, loss: 0.00257413019426167
step: 390, loss: 0.0001634461514186114
step: 400, loss: 0.0013819142477586865
step: 410, loss: 0.0020872089080512524
step: 420, loss: 0.00015487027121707797
step: 430, loss: 2.1915459001320414e-05
step: 440, loss: 0.0002004901471082121
step: 450, loss: 0.00010309371282346547
step: 460, loss: 4.3640724470606074e-05
step: 470, loss: 7.585696585010737e-05
step: 480, loss: 3.9225382352015004e-05
step: 490, loss: 0.0004054027667734772
step: 500, loss: 0.0056878747418522835
step: 510, loss: 0.0003404395538382232
step: 520, loss: 4.210786210023798e-05

```
step: 530, loss: 0.0016519851051270962
/usr/local/lib/python3.7/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=0.9497416627524659, f1=0.946135831381733,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023389568086713552
step: 10, loss: 0.00031514454167336226
step: 20, loss: 0.0001740059960866347
step: 30, loss: 0.0002484697033651173
step: 40, loss: 3.14138233079575e-05
step: 50, loss: 6.700240192003548e-05
step: 60, loss: 0.022667212411761284
step: 70, loss: 0.0008261104812845588
step: 80, loss: 2.6251078452332877e-05
step: 90, loss: 3.198405829607509e-05
step: 100, loss: 0.0022533542942255735
step: 110, loss: 0.00020051577303092927
step: 120, loss: 2.4887856852728873e-05
step: 130, loss: 5.5203050578711554e-05
step: 140, loss: 0.059030674397945404
step: 150, loss: 1.5720504961791448e-05
step: 160, loss: 8.73784301802516e-05
step: 170, loss: 3.517934965202585e-05
step: 180, loss: 0.00018000452837441117
step: 190, loss: 4.276048639439978e-05
step: 200, loss: 0.005245836451649666
step: 210, loss: 0.0002673401904758066
step: 220, loss: 0.0026213235687464476
step: 230, loss: 1.7955328075913712e-05
step: 240, loss: 8.466654981020838e-05
step: 250, loss: 2.001173561438918e-05
step: 260, loss: 2.0808842236874625e-05
step: 270, loss: 4.9249585572397336e-05
step: 280, loss: 4.032019933219999e-05
step: 290, loss: 0.00038631734787486494
step: 300, loss: 8.724098734091967e-05
step: 310, loss: 2.6087376681971364e-05
step: 320, loss: 3.3484710002085194e-05
step: 330, loss: 6.509977538371459e-05
step: 340, loss: 6.427236075978726e-05
step: 350, loss: 0.00012399916886352003
step: 360, loss: 1.4819003808952402e-05
step: 370, loss: 0.00032635018578730524
step: 380, loss: 2.5502209609840065e-05
```

```

step: 390, loss: 1.8782542611006647e-05
step: 400, loss: 1.6327720004483126e-05
step: 410, loss: 2.3822694856789894e-05
step: 420, loss: 2.5744700906216167e-05
step: 430, loss: 2.0622790543711744e-05
step: 440, loss: 0.001205931301228702
step: 450, loss: 0.00011775405437219888
step: 460, loss: 5.671314283972606e-05
step: 470, loss: 0.0001088258795789443
step: 480, loss: 0.0008655881392769516
step: 490, loss: 0.0002011218894040212
step: 500, loss: 3.492090763757005e-05
step: 510, loss: 6.705422856612131e-05
step: 520, loss: 3.191331779817119e-05
step: 530, loss: 2.63734382315306e-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=0.9515828677839852, f1=0.9462465245597776,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.912647975492291e-05
step: 10, loss: 1.8931463273474947e-05
step: 20, loss: 3.520194150041789e-05
step: 30, loss: 0.00011683282355079427
step: 40, loss: 0.00037132666329853237
step: 50, loss: 0.049226514995098114
step: 60, loss: 0.0005634956760331988
step: 70, loss: 2.9648072086274624e-05
step: 80, loss: 2.300292362633627e-05
step: 90, loss: 0.0002562793088145554
step: 100, loss: 2.943141953437589e-05
step: 110, loss: 2.8456370273488574e-05
step: 120, loss: 1.7437436326872557e-05
step: 130, loss: 0.0015351870097219944
step: 140, loss: 2.247777047159616e-05
step: 150, loss: 2.0249955923645757e-05
step: 160, loss: 1.6011081243050285e-05
step: 170, loss: 1.837283343775198e-05
step: 180, loss: 0.0002800055081024766
step: 190, loss: 2.7930736905545928e-05
step: 200, loss: 0.0005034224013797939
step: 210, loss: 2.0082659830222838e-05
step: 220, loss: 0.0001308201754000038
step: 230, loss: 2.3415781470248476e-05
step: 240, loss: 2.2894284484209493e-05

```

```

step: 250, loss: 1.5910494767013006e-05
step: 260, loss: 1.5675812392146327e-05
step: 270, loss: 5.19842142239213e-05
step: 280, loss: 0.0010607248404994607
step: 290, loss: 5.0867776735685766e-05
step: 300, loss: 0.0008977146353572607
step: 310, loss: 3.109283352387138e-05
step: 320, loss: 1.742255153658334e-05
step: 330, loss: 3.781679697567597e-05
step: 340, loss: 1.639474430703558e-05
step: 350, loss: 3.997857857029885e-05
step: 360, loss: 0.00010007378295995295
step: 370, loss: 3.5259250580566004e-05
step: 380, loss: 8.321931818500161e-05
step: 390, loss: 2.6642073862603866e-05
step: 400, loss: 0.00036445262958295643
step: 410, loss: 0.0001634597429074347
step: 420, loss: 0.002783156232908368
step: 430, loss: 0.008390079252421856
step: 440, loss: 1.8965114577440545e-05
step: 450, loss: 6.953886622795835e-05
step: 460, loss: 0.000346797431120649
step: 470, loss: 1.8640714188222773e-05
step: 480, loss: 4.738809002446942e-05
step: 490, loss: 5.6117783969966695e-05
step: 500, loss: 3.0132980100461282e-05
step: 510, loss: 0.001988676842302084
step: 520, loss: 1.6927460819715634e-05
step: 530, loss: 5.376209082896821e-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=0.9482109227871939, f1=0.9454887218045113,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002053932985290885
step: 10, loss: 2.2183123292052187e-05
step: 20, loss: 1.0210927939624526e-05
step: 30, loss: 5.840528319822624e-05
step: 40, loss: 0.0019049086840823293
step: 50, loss: 1.7605418179300614e-05
step: 60, loss: 0.00023461565433535725
step: 70, loss: 1.2986163710593246e-05
step: 80, loss: 1.3694011613551993e-05
step: 90, loss: 0.029545186087489128
step: 100, loss: 0.00020697576110251248

```

```
step: 110, loss: 0.00013223662972450256
step: 120, loss: 3.038463182747364e-05
step: 130, loss: 1.769470691215247e-05
step: 140, loss: 2.2112702936283313e-05
step: 150, loss: 0.002424383070319891
step: 160, loss: 2.1598676539724693e-05
step: 170, loss: 0.00014826253755018115
step: 180, loss: 2.372177186771296e-05
step: 190, loss: 3.67884204024449e-05
step: 200, loss: 0.001846765517257154
step: 210, loss: 3.610217754612677e-05
step: 220, loss: 5.9907262766500935e-05
step: 230, loss: 1.2539197996375151e-05
step: 240, loss: 6.527479126816615e-05
step: 250, loss: 3.883997487719171e-05
step: 260, loss: 1.8599628674564883e-05
step: 270, loss: 2.338896806864068e-05
step: 280, loss: 2.1438307157950476e-05
step: 290, loss: 2.566128023318015e-05
step: 300, loss: 0.00029160871054045856
step: 310, loss: 0.00039085844764485955
step: 320, loss: 0.005274299532175064
step: 330, loss: 0.00160923320800066
step: 340, loss: 2.8245547582628205e-05
step: 350, loss: 1.6268006220343523e-05
step: 360, loss: 1.572797555127181e-05
step: 370, loss: 9.021318692248315e-05
step: 380, loss: 6.923400360392407e-05
step: 390, loss: 0.00040640385122969747
step: 400, loss: 4.3592688598437235e-05
step: 410, loss: 0.0024362981785088778
step: 420, loss: 1.5340516256401315e-05
step: 430, loss: 0.001122022862546146
step: 440, loss: 1.5083405742188916e-05
step: 450, loss: 0.0009626863175071776
step: 460, loss: 0.00861324928700924
step: 470, loss: 1.8617736714077182e-05
step: 480, loss: 1.856956623669248e-05
step: 490, loss: 1.9132667148369364e-05
step: 500, loss: 2.1071509763714857e-05
step: 510, loss: 1.8364942661719397e-05
step: 520, loss: 0.001955340150743723
step: 530, loss: 3.255498450016603e-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=0.9514925373134329, f1=0.9504400185270958,
best_f1=0.9458583988894032
```

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

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

step: 0, loss: 0.00028379185823723674
step: 10, loss: 1.0237017704639584e-05
step: 20, loss: 4.821022957912646e-05
step: 30, loss: 2.1740348529419862e-05
step: 40, loss: 3.284356353105977e-05
step: 50, loss: 8.715331932762638e-05
step: 60, loss: 1.1827621165139135e-05
step: 70, loss: 1.3764775758318137e-05
step: 80, loss: 6.699041841784492e-05
step: 90, loss: 1.3734991625824478e-05
step: 100, loss: 5.0350543460808694e-05
step: 110, loss: 1.8465900211595e-05
step: 120, loss: 2.4272982045658864e-05
step: 130, loss: 0.019464293494820595
step: 140, loss: 0.0008524205186404288
step: 150, loss: 1.3656688679475337e-05
step: 160, loss: 0.0003738909144885838
step: 170, loss: 0.00031464805942960083
step: 180, loss: 1.4785352504986804e-05
step: 190, loss: 7.509311399189755e-05
step: 200, loss: 8.783518569543958e-05
step: 210, loss: 1.3284241504152305e-05
step: 220, loss: 1.33326648210641e-05
step: 230, loss: 2.4656756067997776e-05
step: 240, loss: 1.2282154784770682e-05
step: 250, loss: 1.9855091522913426e-05
step: 260, loss: 0.000308793707517907
step: 270, loss: 4.153417467023246e-05
step: 280, loss: 1.2691938536590897e-05
step: 290, loss: 0.00028245788416825235
step: 300, loss: 3.681630551000126e-05
step: 310, loss: 0.00013220933033153415
step: 320, loss: 0.00042075212695635855
step: 330, loss: 3.2046522392192855e-05
step: 340, loss: 7.821326289558783e-05
step: 350, loss: 0.000508595141582191
step: 360, loss: 7.319836731767282e-05
step: 370, loss: 1.4386900147655979e-05
step: 380, loss: 1.3530103387893178e-05
step: 390, loss: 0.020824965089559555
step: 400, loss: 6.908483919687569e-05
step: 410, loss: 1.7839389329310507e-05
step: 420, loss: 1.5146673831623048e-05
step: 430, loss: 1.6569862054893747e-05
step: 440, loss: 0.00023158953990787268

```

step: 450, loss: 2.4485612811986357e-05
step: 460, loss: 0.00029537416412495077
step: 470, loss: 2.9957591323181987e-05
step: 480, loss: 1.2963865628989879e-05
step: 490, loss: 1.29265308714821e-05
step: 500, loss: 1.1917069969058502e-05
step: 510, loss: 2.9911832825746387e-05
step: 520, loss: 0.00045203277841210365
step: 530, loss: 2.4167773517547175e-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=0.9523364485981308, f1=0.9489795918367347,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.525857351225568e-05
step: 10, loss: 1.9351784430909902e-05
step: 20, loss: 1.8279170035384595e-05
step: 30, loss: 2.3687338398303837e-05
step: 40, loss: 0.08383355289697647
step: 50, loss: 2.58409036177909e-05
step: 60, loss: 9.011412657855544e-06
step: 70, loss: 1.5992407497833483e-05
step: 80, loss: 3.0438601243076846e-05
step: 90, loss: 2.4760603992035612e-05
step: 100, loss: 5.668509766110219e-05
step: 110, loss: 0.0016592281172052026
step: 120, loss: 0.000983183621428907
step: 130, loss: 0.09476422518491745
step: 140, loss: 2.2283276848611422e-05
step: 150, loss: 1.7318487152806483e-05
step: 160, loss: 2.2234560674405657e-05
step: 170, loss: 2.9161599741200916e-05
step: 180, loss: 1.0430708243802655e-05
step: 190, loss: 2.450044485158287e-05
step: 200, loss: 1.1771793651860207e-05
step: 210, loss: 0.0005027626175433397
step: 220, loss: 1.4382960216607898e-05
step: 230, loss: 1.8823262507794425e-05
step: 240, loss: 2.284891343151685e-05
step: 250, loss: 1.6171210518223234e-05
step: 260, loss: 1.586574944667518e-05
step: 270, loss: 1.319110560871195e-05
step: 280, loss: 1.5035035175969824e-05
step: 290, loss: 1.1451424143160693e-05
step: 300, loss: 1.5083298421814106e-05

```



```

step: 310, loss: 4.4421096390578896e-05
step: 320, loss: 0.0005970788770355284
step: 330, loss: 0.0007419384201057255
step: 340, loss: 1.2788776075467467e-05
step: 350, loss: 1.017739123199135e-05
step: 360, loss: 0.0006636937032453716
step: 370, loss: 1.2192694157420192e-05
step: 380, loss: 1.7907108485815115e-05
step: 390, loss: 2.095762647513766e-05
step: 400, loss: 1.5876794350333512e-05
step: 410, loss: 1.5198847904684953e-05
step: 420, loss: 1.4103772628004663e-05
step: 430, loss: 2.0097626475035213e-05
step: 440, loss: 2.0849307475145906e-05
step: 450, loss: 1.112733025365742e-05
step: 460, loss: 0.004836176056414843
step: 470, loss: 0.0025310178752988577
step: 480, loss: 9.540399332763627e-06
step: 490, loss: 1.0967156413244084e-05
step: 500, loss: 2.4963595933513716e-05
step: 510, loss: 1.4848690625512972e-05
step: 520, loss: 1.1019306839443743e-05
step: 530, loss: 0.0005026392755098641
/usr/local/lib/python3.7/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=0.951846657316503, f1=0.9489795918367347,
best_f1=0.9458583988894032
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:29, 191.81it/s]
load_f1 = 0.9562413634269922
real_f1 = 0.9554432705558108
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 184.71it/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: 0.556220531463623
step: 10, loss: 0.3759399652481079
step: 20, loss: 0.3685263395309448
step: 30, loss: 0.3262777626514435
step: 40, loss: 0.17985649406909943
step: 50, loss: 0.39301782846450806
step: 60, loss: 0.21601413190364838
step: 70, loss: 0.1898602545261383
step: 80, loss: 0.21597769856452942
```

```

step: 90, loss: 0.26592034101486206
step: 100, loss: 0.31419986486434937
step: 110, loss: 0.21165865659713745
step: 120, loss: 0.21857288479804993
step: 130, loss: 0.11212141811847687
step: 140, loss: 0.2042391151189804
step: 150, loss: 0.15237806737422943
step: 160, loss: 0.3080075979232788
step: 170, loss: 0.20696988701820374
step: 180, loss: 0.042475972324609756
step: 190, loss: 0.17908482253551483
step: 200, loss: 0.31842395663261414
step: 210, loss: 0.23075710237026215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6977611940298508, f1=0.7212475633528266,
best_f1=0.7212475633528266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10803131759166718
step: 10, loss: 0.21065650880336761
step: 20, loss: 0.1267586350440979
step: 30, loss: 0.08908025175333023
step: 40, loss: 0.16284842789173126
step: 50, loss: 0.13907140493392944
step: 60, loss: 0.35578402876853943
step: 70, loss: 0.12224902957677841
step: 80, loss: 0.13879388570785522
step: 90, loss: 0.12438786774873734
step: 100, loss: 0.012605312280356884
step: 110, loss: 0.07225251197814941
step: 120, loss: 0.16704314947128296
step: 130, loss: 0.00542409485206008
step: 140, loss: 0.2221154421567917
step: 150, loss: 0.18451324105262756
step: 160, loss: 0.16629643738269806
step: 170, loss: 0.10569059103727341
step: 180, loss: 0.09442269057035446
step: 190, loss: 0.04855507239699364
step: 200, loss: 0.05639372766017914
step: 210, loss: 0.13393504917621613
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7137809187279153, f1=0.7302158273381294,
best_f1=0.7302158273381294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06988697499036789
step: 10, loss: 0.12613366544246674
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.029188666492700577
step: 30, loss: 0.10309614986181259
step: 40, loss: 0.06088734790682793
step: 50, loss: 0.04662897065281868
step: 60, loss: 0.05736307054758072
step: 70, loss: 0.052238579839468
step: 80, loss: 0.16900445520877838
step: 90, loss: 0.025891907513141632
step: 100, loss: 0.21878273785114288
step: 110, loss: 0.13302765786647797
step: 120, loss: 0.11671242117881775
step: 130, loss: 0.09286824613809586
step: 140, loss: 0.14360074698925018
step: 150, loss: 0.12259508669376373
step: 160, loss: 0.017391879111528397
step: 170, loss: 0.056820664554834366
step: 180, loss: 0.11872611194849014
step: 190, loss: 0.26301801204681396
step: 200, loss: 0.060905322432518005
step: 210, loss: 0.12833112478256226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7232323232323232, f1=0.7317073170731708,
best_f1=0.7317073170731708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04975586384534836
step: 10, loss: 0.019545026123523712

```

```

step: 20, loss: 0.0604572519659996
step: 30, loss: 0.18798811733722687
step: 40, loss: 0.015379556454718113
step: 50, loss: 0.10644207149744034
step: 60, loss: 0.09579180181026459
step: 70, loss: 0.1693943589925766
step: 80, loss: 0.01879778690636158
step: 90, loss: 0.015107274055480957
step: 100, loss: 0.14396356046199799
step: 110, loss: 0.10562501102685928
step: 120, loss: 0.035410650074481964
step: 130, loss: 0.10542114824056625
step: 140, loss: 0.06943891942501068
step: 150, loss: 0.020542969927191734
step: 160, loss: 0.10885314643383026
step: 170, loss: 0.11490119248628616
step: 180, loss: 0.3177376687526703
step: 190, loss: 0.024639466777443886
step: 200, loss: 0.11049537360668182
step: 210, loss: 0.09281372278928757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7307692307692308, f1=0.7167630057803468,
best_f1=0.7167630057803468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10214332491159439
step: 10, loss: 0.06437429040670395
step: 20, loss: 0.04997345060110092
step: 30, loss: 0.014356770552694798
step: 40, loss: 0.006007056683301926
step: 50, loss: 0.04749506339430809
step: 60, loss: 0.04249594360589981
step: 70, loss: 0.05582045763731003
step: 80, loss: 0.018561027944087982
step: 90, loss: 0.09777853637933731
step: 100, loss: 0.0012027607299387455
step: 110, loss: 0.06051161512732506
step: 120, loss: 0.10757361352443695
step: 130, loss: 0.026290347799658775
step: 140, loss: 0.01376586314290762
step: 150, loss: 0.04056929424405098

```

```

step: 160, loss: 0.12982279062271118
step: 170, loss: 0.050725843757390976
step: 180, loss: 0.04415890574455261
step: 190, loss: 0.044378917664289474
step: 200, loss: 0.06732465326786041
step: 210, loss: 0.02782190591096878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7344632768361581, f1=0.7172675521821633,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012533944100141525
step: 10, loss: 0.022713353857398033
step: 20, loss: 0.008325445465743542
step: 30, loss: 0.0005169732612557709
step: 40, loss: 0.0304974764585495
step: 50, loss: 0.00508063193410635
step: 60, loss: 0.11314231902360916
step: 70, loss: 0.06857115030288696
step: 80, loss: 0.055064693093299866
step: 90, loss: 0.08070835471153259
step: 100, loss: 0.003904033685103059
step: 110, loss: 0.005791744217276573
step: 120, loss: 0.1665387600660324
step: 130, loss: 0.012626505456864834
step: 140, loss: 0.06354540586471558
step: 150, loss: 0.016354534775018692
step: 160, loss: 0.003935771528631449
step: 170, loss: 0.01612585410475731
step: 180, loss: 0.037781838327646255
step: 190, loss: 0.10530197620391846
step: 200, loss: 0.0035569374449551105
step: 210, loss: 0.026287583634257317
/usr/local/lib/python3.7/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=0.7344632768361581, f1=0.7124304267161411,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0038412008434534073
step: 10, loss: 0.06895887106657028
step: 20, loss: 0.02362997643649578
step: 30, loss: 0.002029040362685919
step: 40, loss: 0.0008730938425287604
step: 50, loss: 0.058449387550354004
step: 60, loss: 0.022771380841732025
step: 70, loss: 0.008440670557320118
step: 80, loss: 0.03438124433159828
step: 90, loss: 0.029314761981368065
step: 100, loss: 0.002548872260376811
step: 110, loss: 0.011285602115094662
step: 120, loss: 0.07052242010831833
step: 130, loss: 0.006533680949360132
step: 140, loss: 0.002074790420010686
step: 150, loss: 0.00591886043548584
step: 160, loss: 0.007606960367411375
step: 170, loss: 0.0008916317019611597
step: 180, loss: 0.030799316242337227
step: 190, loss: 0.07149866968393326
step: 200, loss: 0.002915458520874381
step: 210, loss: 0.05374060571193695
/usr/local/lib/python3.7/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=0.7071428571428571, f1=0.7023172905525846,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17187729477882385
step: 10, loss: 0.05918659269809723
step: 20, loss: 0.003782737534493208
step: 30, loss: 0.024737408384680748
step: 40, loss: 0.003974955063313246
step: 50, loss: 0.021364402025938034
step: 60, loss: 0.025634674355387688
step: 70, loss: 0.01633203774690628
step: 80, loss: 0.03034248948097229
step: 90, loss: 0.008715011179447174
step: 100, loss: 0.021365057677030563
step: 110, loss: 0.04438743740320206
step: 120, loss: 0.006998351775109768
step: 130, loss: 0.0004005122173111886
step: 140, loss: 0.026549935340881348
step: 150, loss: 0.00728312274441123
step: 160, loss: 0.02845575474202633
step: 170, loss: 0.0006828156765550375

```



```

step: 180, loss: 0.15520724654197693
step: 190, loss: 0.0415700227022171
step: 200, loss: 0.0014210374793037772
step: 210, loss: 0.08644293993711472
/usr/local/lib/python3.7/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=0.7298387096774194, f1=0.7107438016528925,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007199642714112997
step: 10, loss: 0.006083090323954821
step: 20, loss: 0.0005181028973311186
step: 30, loss: 0.0019778020214289427
step: 40, loss: 0.02099449560046196
step: 50, loss: 0.001067398232407868
step: 60, loss: 0.0822078064084053
step: 70, loss: 0.000533345970325172
step: 80, loss: 0.0016841606702655554
step: 90, loss: 0.0003891613450832665
step: 100, loss: 0.0007888143882155418
step: 110, loss: 0.06173885986208916
step: 120, loss: 0.0010248477337881923
step: 130, loss: 0.006015990860760212
step: 140, loss: 0.010434556752443314
step: 150, loss: 0.07596675306558609
step: 160, loss: 0.0004901195061393082
step: 170, loss: 0.006245005875825882
step: 180, loss: 0.03207165747880936
step: 190, loss: 0.0003273315087426454
step: 200, loss: 0.05063343048095703
step: 210, loss: 0.016147417947649956
/usr/local/lib/python3.7/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=0.7178423236514523, f1=0.7124463519313305,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05192838981747627
step: 10, loss: 0.006121713202446699
step: 20, loss: 0.0003177804755978286
step: 30, loss: 0.024086860939860344
step: 40, loss: 0.0013001230545341969
step: 50, loss: 0.0026112378109246492

```

```

step: 60, loss: 0.024563461542129517
step: 70, loss: 0.04626145586371422
step: 80, loss: 0.00331304594874382
step: 90, loss: 0.1976158320903778
step: 100, loss: 0.018637005239725113
step: 110, loss: 0.0003784849541261792
step: 120, loss: 0.022928671911358833
step: 130, loss: 0.06085760146379471
step: 140, loss: 0.010901857167482376
step: 150, loss: 0.055857520550489426
step: 160, loss: 0.009634447284042835
step: 170, loss: 0.026748033240437508
step: 180, loss: 0.005864458624273539
step: 190, loss: 0.08645538985729218
step: 200, loss: 0.001312682987190783
step: 210, loss: 0.06595741957426071
/usr/local/lib/python3.7/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=0.7012448132780082, f1=0.7148936170212766,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032030604779720306
step: 10, loss: 0.022450704127550125
step: 20, loss: 0.010157773271203041
step: 30, loss: 0.00024080702860374004
step: 40, loss: 0.0141065688803792
step: 50, loss: 0.025923531502485275
step: 60, loss: 0.031177056953310966
step: 70, loss: 0.0044000158086419106
step: 80, loss: 0.1462995558977127
step: 90, loss: 0.02664642408490181
step: 100, loss: 0.01260025892406702
step: 110, loss: 0.026854349300265312
step: 120, loss: 0.0525723397731781
step: 130, loss: 0.01144226174801588
step: 140, loss: 0.007328639272600412
step: 150, loss: 0.0002677045704331249
step: 160, loss: 0.016185728833079338
step: 170, loss: 0.009262275882065296
step: 180, loss: 0.0007849278044886887
step: 190, loss: 0.0002990286739077419
step: 200, loss: 0.0177906583994627
step: 210, loss: 0.0018292745808139443
/usr/local/lib/python3.7/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=0.7227926078028748, f1=0.7172995780590717,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04610787704586983
step: 10, loss: 0.00029478047508746386
step: 20, loss: 0.0110812708735466
step: 30, loss: 0.001755653996951878
step: 40, loss: 0.008906749077141285
step: 50, loss: 0.001132720266468823
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.0009182559442706406
step: 70, loss: 0.0064344992861151695
step: 80, loss: 0.0031464803032577038
step: 90, loss: 0.035545043647289276
step: 100, loss: 0.011313625611364841
step: 110, loss: 0.0005112002836540341
step: 120, loss: 0.0006436094408854842
step: 130, loss: 0.0001523751998320222
step: 140, loss: 0.0002184563927585259
step: 150, loss: 0.00215649395249784
step: 160, loss: 0.004968254826962948
step: 170, loss: 0.0006677487399429083
step: 180, loss: 0.005051173269748688
step: 190, loss: 0.0004175517533440143
step: 200, loss: 0.0003380568523425609
step: 210, loss: 0.02978595346212387
/usr/local/lib/python3.7/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=0.6941431670281997, f1=0.7098214285714286,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011163752526044846
step: 10, loss: 0.00020002669771201909
step: 20, loss: 0.026922941207885742
step: 30, loss: 0.12190567702054977
step: 40, loss: 0.02959483675658703
step: 50, loss: 0.004428020678460598
step: 60, loss: 0.0003552716807462275
step: 70, loss: 0.03410119563341141
step: 80, loss: 0.004286793060600758
step: 90, loss: 0.0002531796635594219
step: 100, loss: 0.003355151042342186

```

```

step: 110, loss: 0.0005887274746783078
step: 120, loss: 0.0003505607892293483
step: 130, loss: 0.0029812271241098642
step: 140, loss: 0.17594285309314728
step: 150, loss: 0.005752273369580507
step: 160, loss: 0.013903530314564705
step: 170, loss: 0.0033337795175611973
step: 180, loss: 0.03337598219513893
step: 190, loss: 0.0008007828728295863
step: 200, loss: 0.0004920808714814484
step: 210, loss: 0.001316581154242158
/usr/local/lib/python3.7/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=0.7113821138211384, f1=0.7219917012448134,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017433850734960288
step: 10, loss: 0.03420064225792885
step: 20, loss: 0.0014858689391985536
step: 30, loss: 0.00015525227354373783
step: 40, loss: 0.0003843284212052822
step: 50, loss: 0.08751172572374344
step: 60, loss: 0.0006726171122863889
step: 70, loss: 0.0018942916067317128
step: 80, loss: 0.00026456190971657634
step: 90, loss: 0.00035711866803467274
step: 100, loss: 0.0008217245922423899
step: 110, loss: 0.000354324554791674
step: 120, loss: 0.039536621421575546
step: 130, loss: 0.0003648936690296978
step: 140, loss: 0.036385804414749146
step: 150, loss: 0.0003163214714732021
step: 160, loss: 0.004313759971410036
step: 170, loss: 0.0001880639756564051
step: 180, loss: 0.0006009121425449848
step: 190, loss: 0.0009603920625522733
step: 200, loss: 0.00022753962548449636
step: 210, loss: 0.0015346803702414036
/usr/local/lib/python3.7/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=0.7051546391752577, f1=0.7172995780590717,
best_f1=0.7172675521821633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002885111316572875
step: 10, loss: 0.0001701623114058748
step: 20, loss: 0.010876242071390152
step: 30, loss: 0.012532033026218414
step: 40, loss: 0.000795372121501714
step: 50, loss: 0.0001689367345534265
step: 60, loss: 0.0005157856503501534
step: 70, loss: 0.0016777943819761276
step: 80, loss: 0.00015922072634566575
step: 90, loss: 0.0189479049295187
step: 100, loss: 0.0007289344212040305
step: 110, loss: 0.00022015803551767021
step: 120, loss: 0.00023326868540607393
step: 130, loss: 0.0005908363382332027
step: 140, loss: 0.00017740592011250556
step: 150, loss: 0.0001410550030414015
step: 160, loss: 0.00025917691527865827
step: 170, loss: 0.00048025575233623385
step: 180, loss: 0.0002405059349257499
step: 190, loss: 0.0005312896682880819
step: 200, loss: 0.0033626705408096313
step: 210, loss: 0.0005320842028595507
/usr/local/lib/python3.7/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=0.7039337474120082, f1=0.7172995780590717,
best_f1=0.7172675521821633
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, 279.95it/s]
load_f1 = 0.7354596622889306
real_f1 = 0.7344632768361581
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, 185.17it/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/
```

```
Downloading: 100% 433/433 [00:00<00:00, 385kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 255kB/s]
Downloading: 100% 440M/440M [00:10<00:00, 40.1MB/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: 0.5885056853294373
step: 10, loss: 0.3623976707458496
step: 20, loss: 0.3010933995246887
step: 30, loss: 0.44807177782058716
step: 40, loss: 0.4249374270439148
step: 50, loss: 0.3111610412597656
```

```

step: 60, loss: 0.2768346071243286
step: 70, loss: 0.24738647043704987
step: 80, loss: 0.2350318282842636
step: 90, loss: 0.2271573543548584
step: 100, loss: 0.2709333598613739
step: 110, loss: 0.39187151193618774
step: 120, loss: 0.0978415384888649
step: 130, loss: 0.12936550378799438
step: 140, loss: 0.03684123978018761
step: 150, loss: 0.1608128845691681
step: 160, loss: 0.142669677734375
step: 170, loss: 0.14894239604473114
step: 180, loss: 0.03139911964535713
step: 190, loss: 0.0868978425860405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7556818181818182, f1=0.8112676056338027,
best_f1=0.8112676056338027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18641594052314758
step: 10, loss: 0.06361959874629974
step: 20, loss: 0.05400201678276062
step: 30, loss: 0.20328865945339203
step: 40, loss: 0.14409108459949493
step: 50, loss: 0.029027758166193962
step: 60, loss: 0.13853007555007935
step: 70, loss: 0.08494941890239716
step: 80, loss: 0.13484710454940796
step: 90, loss: 0.07565037161111832
step: 100, loss: 0.009778869338333607
step: 110, loss: 0.11614995449781418
step: 120, loss: 0.1300220638513565
step: 130, loss: 0.06778372079133987
step: 140, loss: 0.05934769660234451
step: 150, loss: 0.08714204281568527
step: 160, loss: 0.017263520509004593
step: 170, loss: 0.14504781365394592
step: 180, loss: 0.06440846621990204
step: 190, loss: 0.0241332296282053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7906976744186046, f1=0.8056338028169014,
best_f1=0.8056338028169014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022314416244626045
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.28574007749557495
step: 20, loss: 0.022689273580908775
step: 30, loss: 0.02440507896244526
step: 40, loss: 0.11754868179559708
step: 50, loss: 0.08001905679702759
step: 60, loss: 0.03133651986718178
step: 70, loss: 0.1678735762834549
step: 80, loss: 0.07081396877765656
step: 90, loss: 0.07940100878477097
step: 100, loss: 0.020869437605142593
step: 110, loss: 0.005810639355331659
step: 120, loss: 0.23782199621200562
step: 130, loss: 0.03921594098210335
step: 140, loss: 0.11439657211303711
step: 150, loss: 0.20219507813453674
step: 160, loss: 0.021341361105442047
step: 170, loss: 0.04983517527580261
step: 180, loss: 0.06349948048591614
step: 190, loss: 0.13295622169971466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8150134048257373, f1=0.8121546961325967,
best_f1=0.8121546961325967
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011768072843551636
step: 10, loss: 0.01961524970829487
step: 20, loss: 0.05685488507151604
step: 30, loss: 0.02906840294599533
step: 40, loss: 0.05074387416243553

```

```

step: 50, loss: 0.06661883741617203
step: 60, loss: 0.0032518852967768908
step: 70, loss: 0.12546835839748383
step: 80, loss: 0.04096456989645958
step: 90, loss: 0.016084378585219383
step: 100, loss: 0.02605992555618286
step: 110, loss: 0.004486619494855404
step: 120, loss: 0.003979734610766172
step: 130, loss: 0.08574722707271576
step: 140, loss: 0.024633698165416718
step: 150, loss: 0.014722892083227634
step: 160, loss: 0.004228736273944378
step: 170, loss: 0.10113275796175003
step: 180, loss: 0.05781927332282066
step: 190, loss: 0.07732231914997101
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8392370572207084, f1=0.8142076502732241,
best_f1=0.8142076502732241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0677536278963089
step: 10, loss: 0.012589672580361366
step: 20, loss: 0.0019071693532168865
step: 30, loss: 0.009504426270723343
step: 40, loss: 0.15613338351249695
step: 50, loss: 0.13265131413936615
step: 60, loss: 0.039263516664505005
step: 70, loss: 0.04237978160381317
step: 80, loss: 0.03753099590539932
step: 90, loss: 0.010228648781776428
step: 100, loss: 0.0016460224287584424
step: 110, loss: 0.0015592370182275772
step: 120, loss: 0.001986662158742547
step: 130, loss: 0.0907881110906601
step: 140, loss: 0.05013071000576019
step: 150, loss: 0.0041780779138207436
step: 160, loss: 0.051641978323459625
step: 170, loss: 0.006228111684322357
step: 180, loss: 0.11056412756443024
step: 190, loss: 0.009278438054025173
/usr/local/lib/python3.7/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=0.8365650969529086, f1=0.8125, best_f1=0.8142076502732241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05521097779273987
step: 10, loss: 0.04474528878927231
step: 20, loss: 0.005777372047305107
step: 30, loss: 0.001473772106692195
step: 40, loss: 0.07309909164905548
step: 50, loss: 0.03194503113627434
step: 60, loss: 0.062325432896614075
step: 70, loss: 0.0034013011027127504
step: 80, loss: 0.0830734372138977
step: 90, loss: 0.03936631977558136
step: 100, loss: 0.0012260643998160958
step: 110, loss: 0.005187204573303461
step: 120, loss: 0.008175513707101345
step: 130, loss: 0.01052773091942072
step: 140, loss: 0.003315757727250457
step: 150, loss: 0.0024483525194227695
step: 160, loss: 0.010443677194416523
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.10708387941122055
step: 180, loss: 0.06676442176103592
step: 190, loss: 0.007371625397354364
/usr/local/lib/python3.7/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=0.8229166666666667, f1=0.8210526315789474,
best_f1=0.8142076502732241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037931911647319794
step: 10, loss: 0.004887384828180075
step: 20, loss: 0.10571635514497757
step: 30, loss: 0.013445255346596241
step: 40, loss: 0.004602326080203056
step: 50, loss: 0.0084118926897645
step: 60, loss: 0.001514436211436987
step: 70, loss: 0.0958280861377716
step: 80, loss: 0.017567606642842293
step: 90, loss: 0.005634763278067112
step: 100, loss: 0.0003911730891559273
step: 110, loss: 0.09340035915374756
step: 120, loss: 0.0034942447673529387

```

```

step: 130, loss: 0.0024675147142261267
step: 140, loss: 0.006560124922543764
step: 150, loss: 0.0047811889089643955
step: 160, loss: 0.01782502420246601
step: 170, loss: 0.002912758616730571
step: 180, loss: 0.005389583297073841
step: 190, loss: 0.0005220380262471735
/usr/local/lib/python3.7/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=0.8281690140845072, f1=0.8111111111111112,
best_f1=0.8142076502732241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003513782285153866
step: 10, loss: 0.034183964133262634
step: 20, loss: 0.0003452985256444663
step: 30, loss: 0.0033454347867518663
step: 40, loss: 0.0004397415614221245
step: 50, loss: 0.06370808184146881
step: 60, loss: 0.008418816141784191
step: 70, loss: 0.0014920433750376105
step: 80, loss: 0.00011328924301778898
step: 90, loss: 0.0002880820247810334
step: 100, loss: 0.002868334762752056
step: 110, loss: 0.00018567830557003617
step: 120, loss: 0.00019828698714263737
step: 130, loss: 0.0006083464249968529
step: 140, loss: 0.0009180018096230924
step: 150, loss: 0.0005071832565590739
step: 160, loss: 0.002897825790569186
step: 170, loss: 0.0287337563931942
step: 180, loss: 0.013345902785658836
step: 190, loss: 0.1530628204345703
/usr/local/lib/python3.7/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=0.828496042216359, f1=0.8207792207792207,
best_f1=0.8142076502732241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010870123514905572
step: 10, loss: 0.03705490753054619
step: 20, loss: 0.001153611927293241
step: 30, loss: 0.002540492219850421
step: 40, loss: 0.001414875267073512

```

```

step: 50, loss: 0.0006922632455825806
step: 60, loss: 0.004233183804899454
step: 70, loss: 0.0003384923911653459
step: 80, loss: 0.0052333250641822815
step: 90, loss: 0.059208206832408905
step: 100, loss: 0.007927394472062588
step: 110, loss: 0.0011646862840279937
step: 120, loss: 0.022486383095383644
step: 130, loss: 0.012173028662800789
step: 140, loss: 0.0013550767907872796
step: 150, loss: 0.001028150669299066
step: 160, loss: 0.0002551810466684401
step: 170, loss: 0.0007722502923570573
step: 180, loss: 0.0007488872506655753
step: 190, loss: 0.00041571963811293244
/usr/local/lib/python3.7/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=0.8337595907928389, f1=0.8219895287958114,
best_f1=0.8142076502732241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030443596187978983
step: 10, loss: 0.0012444747844710946
step: 20, loss: 0.04985111951828003
step: 30, loss: 0.00020654284162446856
step: 40, loss: 0.023715317249298096
step: 50, loss: 0.00043967048986814916
step: 60, loss: 0.0005345494137145579
step: 70, loss: 0.13819408416748047
step: 80, loss: 0.00023720980971120298
step: 90, loss: 0.00042487040627747774
step: 100, loss: 0.0005242814077064395
step: 110, loss: 0.006259807851165533
step: 120, loss: 0.18404868245124817
step: 130, loss: 0.00040798712871037424
step: 140, loss: 0.0008672262774780393
step: 150, loss: 0.00047982539399527013
step: 160, loss: 0.008879601024091244
step: 170, loss: 0.0003480656014289707
step: 180, loss: 0.0006005446193739772
step: 190, loss: 0.003151924116536975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8421052631578947, f1=0.8263157894736842,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027659302577376366
step: 10, loss: 0.0006149755790829659
step: 20, loss: 0.005958041176199913
step: 30, loss: 0.015240354463458061
step: 40, loss: 0.000407502957386896
step: 50, loss: 0.0006010019569657743
step: 60, loss: 0.0003386044118087739
step: 70, loss: 0.0007643462158739567
step: 80, loss: 0.0013052118010818958
step: 90, loss: 0.0005422004614956677
step: 100, loss: 0.000210541722481139
step: 110, loss: 0.00018622218340169638
step: 120, loss: 0.0002530958445277065
step: 130, loss: 0.0003510559326969087
step: 140, loss: 0.00028587583801709116
step: 150, loss: 0.000156693859025836
step: 160, loss: 0.00041416296153329313
step: 170, loss: 0.00018737306527327746
step: 180, loss: 0.0001729408832034096
step: 190, loss: 0.00028236606158316135
/usr/local/lib/python3.7/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=0.8292682926829268, f1=0.8392370572207084,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011249905946897343
step: 10, loss: 0.0037046868819743395
step: 20, loss: 0.00013509449490811676
step: 30, loss: 0.004643444903194904
step: 40, loss: 0.011738138273358345
step: 50, loss: 0.04039253666996956
step: 60, loss: 0.00036887923488393426
step: 70, loss: 0.00025660908431746066
step: 80, loss: 0.0002524557348806411
step: 90, loss: 0.0001322420284850523
step: 100, loss: 0.00021556804131250829
step: 110, loss: 0.00034102669451385736
step: 120, loss: 0.00013500144996214658

```

```

step: 130, loss: 0.00015546965005341917
step: 140, loss: 0.0005153853562660515
step: 150, loss: 0.003739228006452322
step: 160, loss: 8.228381193475798e-05
step: 170, loss: 0.00038304083864204586
step: 180, loss: 0.00019281971617601812
step: 190, loss: 0.00010100095823872834
/usr/local/lib/python3.7/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=0.8395721925133689, f1=0.8355795148247979,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005161829176358879
step: 10, loss: 0.0005137517000548542
step: 20, loss: 0.00030044346931390464
step: 30, loss: 0.0006218014168553054
step: 40, loss: 0.00016864246572367847
step: 50, loss: 0.0002932421339210123
step: 60, loss: 0.0005170269287191331
step: 70, loss: 0.004433451686054468
step: 80, loss: 0.0001556243369122967
step: 90, loss: 0.001068939920514822
step: 100, loss: 0.0006109887617640197
step: 110, loss: 0.00018721653032116592
step: 120, loss: 0.0010091898730024695
step: 130, loss: 0.00012564606731757522
step: 140, loss: 8.610127406427637e-05
step: 150, loss: 9.526643407298252e-05
step: 160, loss: 0.00011661648022709414
step: 170, loss: 0.0002577351115178317
step: 180, loss: 0.0001373546983813867
step: 190, loss: 0.0001958595385076478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8480000000000001, f1=0.8342245989304813,
best_f1=0.8342245989304813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001366545184282586

```

```

step: 10, loss: 9.955611312761903e-05
step: 20, loss: 0.00023162477009464055
step: 30, loss: 9.4264090876095e-05
step: 40, loss: 0.00019340545986779034
step: 50, loss: 0.0001804835192160681
step: 60, loss: 0.00046048115473240614
step: 70, loss: 9.352310735266656e-05
step: 80, loss: 8.86696288944222e-05
step: 90, loss: 0.00976009014993906
step: 100, loss: 0.002391789574176073
step: 110, loss: 0.00014109529729466885
step: 120, loss: 0.0004018909530714154
step: 130, loss: 0.00012066858471371233
step: 140, loss: 0.0009877310367301106
step: 150, loss: 0.00011882076069014147
step: 160, loss: 0.00019061347120441496
step: 170, loss: 0.0002740056370384991
step: 180, loss: 0.0016685285372659564
step: 190, loss: 0.00029611485661007464
/usr/local/lib/python3.7/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=0.8395721925133689, f1=0.8292682926829268,
best_f1=0.8342245989304813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006156439194455743
step: 10, loss: 0.00011307634849799797
step: 20, loss: 0.00010142418614123017
step: 30, loss: 0.00010520398063817993
step: 40, loss: 9.585201769368723e-05
step: 50, loss: 0.0002667894877959043
step: 60, loss: 0.0002670838439371437
step: 70, loss: 0.011904286220669746
step: 80, loss: 0.00034887235960923135
step: 90, loss: 6.847163604106754e-05
step: 100, loss: 0.0001870826235972345
step: 110, loss: 0.0003418915730435401
step: 120, loss: 0.0001363271294394508
step: 130, loss: 9.785551810637116e-05
step: 140, loss: 0.0002075265219900757
step: 150, loss: 0.0001504352258052677
step: 160, loss: 8.995657844934613e-05
step: 170, loss: 9.514450357528403e-05
step: 180, loss: 0.00032917968928813934
step: 190, loss: 0.00010769462824100628
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.8486486486486486, f1=0.8360655737704918,
best_f1=0.8360655737704918
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 [00:11, 171.50it/s]
load_f1 = 0.8509485094850948
real_f1 = 0.842391304347826
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, 182.59it/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 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: 0.6254998445510864
step: 10, loss: 0.35666558146476746
step: 20, loss: 0.314370334148407
step: 30, loss: 0.3767809569835663
step: 40, loss: 0.2725193202495575
step: 50, loss: 0.24075987935066223
step: 60, loss: 0.24154114723205566
step: 70, loss: 0.33150288462638855
step: 80, loss: 0.2887968420982361
step: 90, loss: 0.17655086517333984
step: 100, loss: 0.28984367847442627
step: 110, loss: 0.19561214745044708
step: 120, loss: 0.10848009586334229
step: 130, loss: 0.0488961786031723
step: 140, loss: 0.25948071479797363
step: 150, loss: 0.3182094693183899
step: 160, loss: 0.12827181816101074
step: 170, loss: 0.216439351439476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7808219178082191, f1=0.7668161434977578,
best_f1=0.7668161434977578

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04229458421468735
step: 10, loss: 0.10415011644363403
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.021432317793369293
step: 30, loss: 0.2998320162296295
step: 40, loss: 0.04928302392363548
step: 50, loss: 0.06742287427186966
step: 60, loss: 0.054702263325452805
step: 70, loss: 0.12573403120040894
step: 80, loss: 0.05205385014414787
step: 90, loss: 0.08195486664772034
step: 100, loss: 0.15191909670829773
step: 110, loss: 0.05833401903510094
step: 120, loss: 0.02377202920615673
step: 130, loss: 0.01943143643438816
step: 140, loss: 0.43569549918174744
step: 150, loss: 0.16460154950618744
step: 160, loss: 0.1089448630809784
step: 170, loss: 0.1685229241847992
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8057553956834533, f1=0.8130841121495326,
best_f1=0.8130841121495326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08929470926523209
step: 10, loss: 0.06801876425743103
step: 20, loss: 0.014277901500463486
step: 30, loss: 0.05356553941965103
step: 40, loss: 0.08680401742458344
step: 50, loss: 0.20583154261112213
step: 60, loss: 0.046110183000564575
step: 70, loss: 0.0202907994389534
step: 80, loss: 0.048947080969810486
step: 90, loss: 0.022093988955020905
step: 100, loss: 0.11204195767641068
step: 110, loss: 0.09401784837245941
step: 120, loss: 0.07818230986595154
step: 130, loss: 0.215781107544899

```

```

step: 140, loss: 0.13590747117996216
step: 150, loss: 0.010641870088875294
step: 160, loss: 0.010006586089730263
step: 170, loss: 0.0359475314617157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8138957816377173, f1=0.8188235294117647,
best_f1=0.8188235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011088519357144833
step: 10, loss: 0.022095942869782448
step: 20, loss: 0.0036710079293698072
step: 30, loss: 0.035695306956768036
step: 40, loss: 0.003867109538987279
step: 50, loss: 0.05522169545292854
step: 60, loss: 0.19865019619464874
step: 70, loss: 0.007983177900314331
step: 80, loss: 0.11580240726470947
step: 90, loss: 0.08519107848405838
step: 100, loss: 0.20804719626903534
step: 110, loss: 0.07053491473197937
step: 120, loss: 0.002140411175787449
step: 130, loss: 0.009605273604393005
step: 140, loss: 0.04643365740776062
step: 150, loss: 0.18875424563884735
step: 160, loss: 0.055929120630025864
step: 170, loss: 0.0810643658041954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8170426065162907, f1=0.826923076923077,
best_f1=0.826923076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010635113343596458
step: 10, loss: 0.010900909081101418

```

```

step: 20, loss: 0.010216943919658661
step: 30, loss: 0.013574928976595402
step: 40, loss: 0.1614471822977066
step: 50, loss: 0.003956209868192673
step: 60, loss: 0.004208070691674948
step: 70, loss: 0.1612349897623062
step: 80, loss: 0.02617041766643524
step: 90, loss: 0.035540029406547546
step: 100, loss: 0.022843526676297188
step: 110, loss: 0.0072678630240261555
step: 120, loss: 0.0077276285737752914
step: 130, loss: 0.020204421132802963
step: 140, loss: 0.08962595462799072
step: 150, loss: 0.027072669938206673
step: 160, loss: 0.04516759514808655
step: 170, loss: 0.026278134435415268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8413793103448276, f1=0.8133333333333335,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006946094799786806
step: 10, loss: 0.006570510100573301
step: 20, loss: 0.003564022947102785
step: 30, loss: 0.006325113121420145
step: 40, loss: 0.0013061959762126207
step: 50, loss: 0.11420997977256775
step: 60, loss: 0.008177450858056545
step: 70, loss: 0.0818186029791832
step: 80, loss: 0.010252335108816624
step: 90, loss: 0.013595778495073318
step: 100, loss: 0.008141608908772469
step: 110, loss: 0.0029915638733655214
step: 120, loss: 0.01544820237904787
step: 130, loss: 0.003204071195796132
step: 140, loss: 0.007852856069803238
step: 150, loss: 0.04338131472468376
step: 160, loss: 0.15195424854755402
step: 170, loss: 0.0026865783147513866
/usr/local/lib/python3.7/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=0.8407310704960835, f1=0.8329048843187661,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0069120293483138084
step: 10, loss: 0.0018850701162591577
step: 20, loss: 0.0028637468349188566
step: 30, loss: 0.021287254989147186
step: 40, loss: 0.0028091324493288994
step: 50, loss: 0.052907779812812805
step: 60, loss: 0.0002187474601669237
step: 70, loss: 0.0009383850265294313
step: 80, loss: 0.02043456956744194
step: 90, loss: 0.00043799614650197327
step: 100, loss: 0.005006765481084585
step: 110, loss: 0.006822560913860798
step: 120, loss: 0.0046430048532783985
step: 130, loss: 0.12030325829982758
step: 140, loss: 0.00037344038719311357
step: 150, loss: 0.001118116662837565
step: 160, loss: 0.0011756758904084563
step: 170, loss: 0.004947534762322903
/usr/local/lib/python3.7/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=0.8280871670702179, f1=0.8177570093457943,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003346348414197564
step: 10, loss: 0.0006855450919829309
step: 20, loss: 0.0004565291164908558
step: 30, loss: 0.030740905553102493
step: 40, loss: 0.0007527315756306052
step: 50, loss: 0.001993655227124691
step: 60, loss: 0.002110763918608427
step: 70, loss: 0.003714231541380286
step: 80, loss: 0.016403233632445335
step: 90, loss: 0.0005321413627825677
step: 100, loss: 0.03092549927532673
step: 110, loss: 0.16712301969528198
step: 120, loss: 0.0002903171116486192
step: 130, loss: 0.0013311353977769613
step: 140, loss: 0.03601919114589691
step: 150, loss: 0.0028253537602722645

```

```

step: 160, loss: 0.0006810212507843971
step: 170, loss: 0.013872732408344746
/usr/local/lib/python3.7/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=0.819047619047619, f1=0.7981651376146789,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001509562716819346
step: 10, loss: 0.0026099528186023235
step: 20, loss: 0.008771151304244995
step: 30, loss: 0.009968549944460392
step: 40, loss: 0.0005824944819323719
step: 50, loss: 0.0002305624366272241
step: 60, loss: 0.005167584866285324
step: 70, loss: 0.007186914328485727
step: 80, loss: 0.002643692772835493
step: 90, loss: 0.015251866541802883
step: 100, loss: 0.0035141126718372107
step: 110, loss: 0.0009453381062485278
step: 120, loss: 0.006935469806194305
step: 130, loss: 0.03416703641414642
step: 140, loss: 0.0002028545568464324
step: 150, loss: 0.0010820769239217043
step: 160, loss: 0.0008117327815853059
step: 170, loss: 0.03517821058630943
/usr/local/lib/python3.7/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=0.8407960199004973, f1=0.8463356973995272,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03496107459068298
step: 10, loss: 0.0022832718677818775
step: 20, loss: 0.2260214388370514
step: 30, loss: 0.012588834390044212
step: 40, loss: 0.0004905920941382647
step: 50, loss: 0.027206484228372574
step: 60, loss: 0.00014950078912079334
step: 70, loss: 0.005935133900493383
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.012712378054857254
step: 90, loss: 0.03519150987267494
step: 100, loss: 0.0070070079527795315

```



```

step: 110, loss: 0.00103346211835742
step: 120, loss: 0.0013575704069808125
step: 130, loss: 0.0004184111312497407
step: 140, loss: 0.000515900319442153
step: 150, loss: 0.0016824720660224557
step: 160, loss: 0.0005792578449472785
step: 170, loss: 0.0002583466994110495
/usr/local/lib/python3.7/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=0.828235294117647, f1=0.8177777777777777,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028501536697149277
step: 10, loss: 0.015040833503007889
step: 20, loss: 0.00023235820117406547
step: 30, loss: 0.00021015683887526393
step: 40, loss: 0.0005148741765879095
step: 50, loss: 0.0003498699516057968
step: 60, loss: 0.0029205544851720333
step: 70, loss: 0.00012932674144394696
step: 80, loss: 0.00013411724648904055
step: 90, loss: 0.0004912407603114843
step: 100, loss: 0.0031819825526326895
step: 110, loss: 0.03582391142845154
step: 120, loss: 0.00019095161405857652
step: 130, loss: 0.00017694228154141456
step: 140, loss: 0.0002183797478210181
step: 150, loss: 0.004229937680065632
step: 160, loss: 0.05290990322828293
step: 170, loss: 0.0003073419793508947
/usr/local/lib/python3.7/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=0.837696335078534, f1=0.827930174563591,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035774875432252884
step: 10, loss: 0.000521876965649426
step: 20, loss: 0.00015498542052228004
step: 30, loss: 0.00015782078844495118
step: 40, loss: 0.00031438402947969735
step: 50, loss: 0.00014432381431106478
step: 60, loss: 0.005698562599718571

```

```

step: 70, loss: 0.09358122199773788
step: 80, loss: 6.950582610443234e-05
step: 90, loss: 0.0002640371094457805
step: 100, loss: 0.00045507983304560184
step: 110, loss: 0.00018693192396312952
step: 120, loss: 0.04570094123482704
step: 130, loss: 0.01691398024559021
step: 140, loss: 0.0137722073122859
step: 150, loss: 0.013385145924985409
step: 160, loss: 0.00019282361608929932
step: 170, loss: 0.00011656076094368473
/usr/local/lib/python3.7/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=0.8369829683698297, f1=0.8271028037383177,
best_f1=0.8133333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018554363399744034
step: 10, loss: 0.000300115963909775
step: 20, loss: 0.0001487774570705369
step: 30, loss: 0.0002786295663099736
step: 40, loss: 0.0005038535455241799
step: 50, loss: 0.00010577194188954309
step: 60, loss: 0.005377796944230795
step: 70, loss: 0.0004915317404083908
step: 80, loss: 0.005446907132863998
step: 90, loss: 9.358136594528332e-05
step: 100, loss: 0.0004056362377014011
step: 110, loss: 7.761317101540044e-05
step: 120, loss: 0.0218641497194767
step: 130, loss: 0.00011348469706717879
step: 140, loss: 0.0022257715463638306
step: 150, loss: 0.09391739964485168
step: 160, loss: 0.009332283399999142
step: 170, loss: 0.006108648143708706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8433734939759036, f1=0.8248847926267281,
best_f1=0.8248847926267281
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020989318727515638
step: 10, loss: 9.948264778358862e-05
step: 20, loss: 0.0002509743790142238
step: 30, loss: 0.0001438693725503981
step: 40, loss: 0.00012820497795473784
step: 50, loss: 0.0001700620341580361
step: 60, loss: 6.517017027363181e-05
step: 70, loss: 0.0001132218458224088
step: 80, loss: 0.0003776235389523208
step: 90, loss: 0.0001541936071589589
step: 100, loss: 7.983107934705913e-05
step: 110, loss: 0.00010572385508567095
step: 120, loss: 0.07339993864297867
step: 130, loss: 0.013408331200480461
step: 140, loss: 0.00017753693100530654
step: 150, loss: 6.718430813634768e-05
step: 160, loss: 5.212253381614573e-05
step: 170, loss: 0.1081944927573204
/usr/local/lib/python3.7/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=0.8382352941176471, f1=0.812206572769953,
best_f1=0.8248847926267281
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013101166405249387
step: 10, loss: 0.0005253115086816251
step: 20, loss: 0.0017721406184136868
step: 30, loss: 8.567313489038497e-05
step: 40, loss: 0.0001475252356613055
step: 50, loss: 0.000237153290072456
step: 60, loss: 0.0009111182298511267
step: 70, loss: 6.437692354666069e-05
step: 80, loss: 0.01698790304362774
step: 90, loss: 0.000700472854077816
step: 100, loss: 0.0004347975773271173
step: 110, loss: 0.000155282934429124
step: 120, loss: 0.0003714092308655381
step: 130, loss: 0.0011557737598195672
step: 140, loss: 0.00029703948530368507
step: 150, loss: 0.0033739274367690086
step: 160, loss: 8.81947489688173e-05
step: 170, loss: 0.00036578511935658753
/usr/local/lib/python3.7/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=0.8353808353808354, f1=0.812206572769953,
best_f1=0.8248847926267281
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 [00:07, 245.74it/s]
```

```
load_f1 = 0.8461538461538461
```

```
real_f1 = 0.8441247002398082
```

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

```
4398it [00:19, 229.15it/s]
```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6281794309616089
step: 10, loss: 0.6173593401908875
step: 20, loss: 0.4149489998817444
step: 30, loss: 0.13769061863422394
step: 40, loss: 0.17023976147174835
step: 50, loss: 0.09668057411909103
step: 60, loss: 0.09598471224308014
step: 70, loss: 0.0563083179295063
step: 80, loss: 0.09121453016996384
step: 90, loss: 0.15916582942008972
step: 100, loss: 0.003994429484009743
step: 110, loss: 0.22906096279621124
step: 120, loss: 0.008732009679079056
step: 130, loss: 0.010687828995287418
step: 140, loss: 0.0015193346189334989
step: 150, loss: 0.01260019838809967
step: 160, loss: 0.015341510064899921
step: 170, loss: 0.13917550444602966
step: 180, loss: 0.0036968139465898275
step: 190, loss: 0.07530252635478973
step: 200, loss: 0.05592353641986847
step: 210, loss: 0.004235477186739445
step: 220, loss: 0.010676039382815361
step: 230, loss: 0.020160987973213196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9755555555555556, 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: 0.008177296258509159
step: 10, loss: 0.002384732011705637
step: 20, loss: 0.0897141620516777
step: 30, loss: 0.1766941100358963
step: 40, loss: 0.034994397312402725
step: 50, loss: 0.005274391267448664
step: 60, loss: 0.003807745175436139
step: 70, loss: 0.08182729780673981
step: 80, loss: 0.002676707459613681
step: 90, loss: 0.0059972964227199554
step: 100, loss: 0.0971975326538086
step: 110, loss: 0.13426735997200012
step: 120, loss: 0.07943293452262878
step: 130, loss: 0.0060943481512367725
step: 140, loss: 0.004393472336232662
step: 150, loss: 0.008509095758199692
step: 160, loss: 0.012829158455133438
step: 170, loss: 0.001133456127718091
step: 180, loss: 0.005112954415380955
step: 190, loss: 0.002580092055723071
step: 200, loss: 0.0067705875262618065
step: 210, loss: 0.0010652871569618583
step: 220, loss: 0.2707439064979553
step: 230, loss: 0.03810650110244751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9821428571428571, f1=0.9733333333333333,
best_f1=0.9733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026948807761073112
step: 10, loss: 0.0036605626810342073
step: 20, loss: 0.002611722331494093

```

```

step: 30, loss: 0.007497904822230339
step: 40, loss: 0.09892063587903976
step: 50, loss: 0.03433392941951752
step: 60, loss: 0.002162155695259571
step: 70, loss: 0.0034405940677970648
step: 80, loss: 0.0006429718341678381
step: 90, loss: 0.016583042219281197
step: 100, loss: 0.0005359880160540342
step: 110, loss: 0.0004748079809360206
step: 120, loss: 0.029060227796435356
step: 130, loss: 0.0017914391355589032
step: 140, loss: 0.003428006311878562
step: 150, loss: 0.09430105239152908
step: 160, loss: 0.028203332796692848
step: 170, loss: 0.004611265379935503
step: 180, loss: 0.0014055435312911868
step: 190, loss: 0.0030956913251429796
step: 200, loss: 0.000805095536634326
step: 210, loss: 0.015157217159867287
step: 220, loss: 0.000812264159321785
step: 230, loss: 0.07318328320980072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832026875699889, f1=0.9788182831661093,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011163774179294705
step: 10, loss: 0.0806613564491272
step: 20, loss: 0.0007323484751395881
step: 30, loss: 0.0008516973466612399
step: 40, loss: 0.008380638435482979
step: 50, loss: 0.0007139304070733488
step: 60, loss: 0.0010695678647607565
step: 70, loss: 0.001455488963983953
step: 80, loss: 0.013947748579084873
step: 90, loss: 0.0014026762219145894
step: 100, loss: 0.0007001115591265261
step: 110, loss: 0.0008402737439610064
step: 120, loss: 0.011046184226870537
step: 130, loss: 0.0005277203745208681
step: 140, loss: 0.0004692124202847481

```



```

step: 150, loss: 0.147408589720726
step: 160, loss: 0.010434574447572231
step: 170, loss: 0.010291491635143757
step: 180, loss: 0.0007706038886681199
step: 190, loss: 0.0028613568283617496
step: 200, loss: 0.004651875235140324
step: 210, loss: 0.17883484065532684
step: 220, loss: 0.0006038487190380692
step: 230, loss: 0.003594489535316825
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9786276715410572,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00125881505664438
step: 10, loss: 0.0011916945222765207
step: 20, loss: 0.0015254486352205276
step: 30, loss: 0.0004960931255482137
step: 40, loss: 0.0007833693525753915
step: 50, loss: 0.01586403138935566
step: 60, loss: 0.015457931905984879
step: 70, loss: 0.0015349223976954818
step: 80, loss: 0.0003581796772778034
step: 90, loss: 0.0009481972083449364
step: 100, loss: 0.001235560397617519
step: 110, loss: 0.0006559640169143677
step: 120, loss: 0.0001387531083310023
step: 130, loss: 0.0003853331145364791
step: 140, loss: 0.0018759971717372537
step: 150, loss: 0.004783699754625559
step: 160, loss: 0.00027122345636598766
step: 170, loss: 0.0017715840367600322
step: 180, loss: 0.001819814438931644
step: 190, loss: 0.07150456309318542
step: 200, loss: 0.0061757490038871765
step: 210, loss: 0.0007541737868450582
step: 220, loss: 0.0007930738502182066
step: 230, loss: 0.0002628942602314055
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 5: dev_f1=0.9865470852017937, f1=0.9853768278965129,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006844268646091223
step: 10, loss: 0.002468990860506892
step: 20, loss: 0.001114921527914703
step: 30, loss: 0.0002052587369689718
step: 40, loss: 0.00021240673959255219
step: 50, loss: 0.0004059630155097693
step: 60, loss: 0.0001565287384437397
step: 70, loss: 0.0004190433246549219
step: 80, loss: 0.0052931541576981544
step: 90, loss: 0.00037003919715061784
step: 100, loss: 0.02548391744494438
step: 110, loss: 0.0002613765827845782
step: 120, loss: 0.006895760539919138
step: 130, loss: 0.0006441084551624954
step: 140, loss: 0.0005695342551916838
step: 150, loss: 0.0012669629650190473
step: 160, loss: 0.0002907465968746692
step: 170, loss: 0.00012581259943544865
step: 180, loss: 0.022729044780135155
step: 190, loss: 0.014756144024431705
step: 200, loss: 9.410150232724845e-05
step: 210, loss: 0.0001555058261146769
step: 220, loss: 0.0002130679931724444
step: 230, loss: 0.05171060562133789
/usr/local/lib/python3.7/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=0.9844097995545658, f1=0.9799554565701558,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033190255635418
step: 10, loss: 8.907931623980403e-05
step: 20, loss: 0.0002238597662653774
step: 30, loss: 0.0009649202693253756
step: 40, loss: 0.00015071866801008582
step: 50, loss: 0.00017223406757693738
step: 60, loss: 0.0031883493065834045
step: 70, loss: 0.025218317285180092
step: 80, loss: 0.00016354442050214857
step: 90, loss: 0.00021494798420462757
step: 100, loss: 8.444749983027577e-05

```

```

step: 110, loss: 0.0001424876827513799
step: 120, loss: 6.593031866941601e-05
step: 130, loss: 5.8336969232186675e-05
step: 140, loss: 6.937118450878188e-05
step: 150, loss: 0.01065848208963871
step: 160, loss: 0.041535843163728714
step: 170, loss: 0.00028302351711317897
step: 180, loss: 0.0004087071865797043
step: 190, loss: 0.00013842985208611935
step: 200, loss: 0.08298049122095108
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 0.00013278608093969524
step: 220, loss: 0.001203457242809236
step: 230, loss: 0.014493327587842941
/usr/local/lib/python3.7/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=0.9776785714285714, f1=0.9732739420935412,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014964252477511764
step: 10, loss: 0.0007964426768012345
step: 20, loss: 0.00012007517216261476
step: 30, loss: 0.00011602832091739401
step: 40, loss: 0.00043296380317769945
step: 50, loss: 0.0003446594055276364
step: 60, loss: 7.53247004467994e-05
step: 70, loss: 0.0001576197100803256
step: 80, loss: 0.001397899934090674
step: 90, loss: 8.045851427596062e-05
step: 100, loss: 0.0001486835681134835
step: 110, loss: 0.06589843332767487
step: 120, loss: 0.10345034301280975
step: 130, loss: 0.0014432445168495178
step: 140, loss: 0.00021571459365077317
step: 150, loss: 0.0001137529470724985
step: 160, loss: 0.00014865215052850544
step: 170, loss: 0.00012663318193517625
step: 180, loss: 0.0008671424584463239
step: 190, loss: 0.023788198828697205
step: 200, loss: 0.016438480466604233
step: 210, loss: 0.040760159492492676
step: 220, loss: 0.0002851762983482331
step: 230, loss: 0.0006966927903704345
/usr/local/lib/python3.7/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=0.9832402234636871, f1=0.9776286353467561,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017021194798871875
step: 10, loss: 0.1111704558134079
step: 20, loss: 0.013808061368763447
step: 30, loss: 0.030414316803216934
step: 40, loss: 0.004140161443501711
step: 50, loss: 8.448890730505809e-05
step: 60, loss: 0.00010076205944642425
step: 70, loss: 9.136383596342057e-05
step: 80, loss: 0.00012328586308285594
step: 90, loss: 0.003016378963366151
step: 100, loss: 0.0009189133415929973
step: 110, loss: 0.00012090452219126746
step: 120, loss: 0.00014846764679532498
step: 130, loss: 0.00014680666208732873
step: 140, loss: 0.00017411456792615354
step: 150, loss: 0.0005588079220615327
step: 160, loss: 0.00014058142551220953
step: 170, loss: 0.00017347559332847595
step: 180, loss: 0.0025925147347152233
step: 190, loss: 0.00013150955783203244
step: 200, loss: 0.01836399734020233
step: 210, loss: 0.00010781851597130299
step: 220, loss: 0.00010874387953663245
step: 230, loss: 9.72452326095663e-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=0.9854423292273236, f1=0.9831271091113611,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007493185694329441
step: 10, loss: 5.861676254426129e-05
step: 20, loss: 0.00025462728808633983
step: 30, loss: 8.044344576774165e-05
step: 40, loss: 0.00016519772179890424
step: 50, loss: 8.244643686339259e-05
step: 60, loss: 0.001821925980038941
step: 70, loss: 0.0004477622569538653
step: 80, loss: 9.200275962939486e-05
step: 90, loss: 0.000248106662184

```

```

step: 100, loss: 0.004827427212148905
step: 110, loss: 0.00030157569563016295
step: 120, loss: 0.0017395700560882688
step: 130, loss: 0.0001832895795814693
step: 140, loss: 0.04655195400118828
step: 150, loss: 0.013339569792151451
step: 160, loss: 0.00014549566549248993
step: 170, loss: 9.268528810935095e-05
step: 180, loss: 0.00014267832739278674
step: 190, loss: 0.008899242617189884
step: 200, loss: 7.417322194669396e-05
step: 210, loss: 0.0003285653656348586
step: 220, loss: 7.483494118787348e-05
step: 230, loss: 0.0012735676718875766
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9854096520763187,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.325395250110887e-05
step: 10, loss: 9.290609887102619e-05
step: 20, loss: 0.0002094213996315375
step: 30, loss: 0.02464340813457966
step: 40, loss: 0.00015554387937299907
step: 50, loss: 7.333978282986209e-05
step: 60, loss: 9.63663260336034e-05
step: 70, loss: 6.791617488488555e-05
step: 80, loss: 4.3338131945347413e-05
step: 90, loss: 0.0003071083046961576
step: 100, loss: 5.6441316701238975e-05
step: 110, loss: 0.020446090027689934
step: 120, loss: 5.362899537431076e-05
step: 130, loss: 4.7011912101879716e-05
step: 140, loss: 6.366043089656159e-05
step: 150, loss: 0.027460604906082153
step: 160, loss: 4.4812590203946456e-05
step: 170, loss: 0.02206052467226982
step: 180, loss: 7.543393439846113e-05
step: 190, loss: 6.233326712390408e-05
step: 200, loss: 5.740371489082463e-05
step: 210, loss: 5.183440225664526e-05
step: 220, loss: 5.0493519665906206e-05
step: 230, loss: 4.679813719121739e-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=0.9854748603351955, f1=0.9788182831661093,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.79368749842979e-05
step: 10, loss: 0.000114729831693694
step: 20, loss: 5.320241689332761e-05
step: 30, loss: 0.00011226241622352973
step: 40, loss: 6.751228647772223e-05
step: 50, loss: 6.489180668722838e-05
step: 60, loss: 0.00016484649677295238
step: 70, loss: 9.016793046612293e-05
step: 80, loss: 6.694212788715959e-05
step: 90, loss: 4.0495920984540135e-05
step: 100, loss: 4.321063534007408e-05
step: 110, loss: 5.948996476945467e-05
step: 120, loss: 5.126958785695024e-05
step: 130, loss: 4.025776797789149e-05
step: 140, loss: 0.014990575611591339
step: 150, loss: 0.00027809542370960116
step: 160, loss: 8.080470433924347e-05
step: 170, loss: 4.892149445367977e-05
step: 180, loss: 0.019864508882164955
step: 190, loss: 0.00017288138042204082
step: 200, loss: 2.927613604697399e-05
step: 210, loss: 7.815352728357539e-05
step: 220, loss: 4.214411092107184e-05
step: 230, loss: 0.05277992784976959
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865771812080537, 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: 0.00012936429993715137
step: 10, loss: 0.00027714730822481215
step: 20, loss: 7.467273826478049e-05
step: 30, loss: 5.5562897614436224e-05
step: 40, loss: 3.4521206544013694e-05
step: 50, loss: 0.00038165325531736016

```

```

step: 60, loss: 4.683536826632917e-05
step: 70, loss: 3.6587029171641916e-05
step: 80, loss: 3.849938002531417e-05
step: 90, loss: 5.1034283387707546e-05
step: 100, loss: 3.5515120544005185e-05
step: 110, loss: 0.018145641312003136
step: 120, loss: 0.022450895980000496
step: 130, loss: 4.653347787098028e-05
step: 140, loss: 5.1941849960712716e-05
step: 150, loss: 4.7796689614187926e-05
step: 160, loss: 4.915874160360545e-05
step: 170, loss: 4.606756556313485e-05
step: 180, loss: 3.7575628084596246e-05
step: 190, loss: 3.276263305451721e-05
step: 200, loss: 6.709710578434169e-05
step: 210, loss: 2.2582269593840465e-05
step: 220, loss: 0.0002292769931955263
step: 230, loss: 6.00126113567967e-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=0.9855072463768116, f1=0.9799554565701558,
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: 9.239806240657344e-05
step: 10, loss: 0.0003401884459890425
step: 20, loss: 3.475937046459876e-05
step: 30, loss: 4.013115176348947e-05
step: 40, loss: 0.014132983051240444
step: 50, loss: 5.161299486644566e-05
step: 60, loss: 4.079692007508129e-05
step: 70, loss: 4.602325861924328e-05
step: 80, loss: 6.174900045152754e-05
step: 90, loss: 7.198968523880467e-05
step: 100, loss: 4.0574064769316465e-05
step: 110, loss: 8.82833992363885e-05
step: 120, loss: 2.5644199922680855e-05
step: 130, loss: 3.624965393100865e-05
step: 140, loss: 0.00036728082341142
step: 150, loss: 3.351527630002238e-05
step: 160, loss: 0.00042563225724734366
step: 170, loss: 5.6724420574028045e-05
step: 180, loss: 6.033379031578079e-05
step: 190, loss: 0.00012685646652244031
step: 200, loss: 2.9425125831039622e-05
step: 210, loss: 3.8562691770493984e-05

```

```

step: 220, loss: 3.3584528864594176e-05
step: 230, loss: 0.0008750489796511829
/usr/local/lib/python3.7/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=0.985539488320356, f1=0.9777777777777777,
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: 2.4619912437628955e-05
step: 10, loss: 2.151681292161811e-05
step: 20, loss: 3.312799526611343e-05
step: 30, loss: 4.587423609336838e-05
step: 40, loss: 0.0005386738339439034
step: 50, loss: 0.027244264259934425
step: 60, loss: 9.318064985563979e-05
step: 70, loss: 5.0683269364526495e-05
step: 80, loss: 3.0483100999845192e-05
step: 90, loss: 7.984867988852784e-05
step: 100, loss: 3.6122819437878206e-05
step: 110, loss: 3.8068872527219355e-05
step: 120, loss: 4.7323501348728314e-05
step: 130, loss: 0.004356815945357084
step: 140, loss: 3.2569274480920285e-05
step: 150, loss: 0.00022993939637672156
step: 160, loss: 2.9019114663242362e-05
step: 170, loss: 3.04271379718557e-05
step: 180, loss: 9.59406970650889e-05
step: 190, loss: 4.407117739901878e-05
step: 200, loss: 3.790670962189324e-05
step: 210, loss: 7.239523256430402e-05
step: 220, loss: 0.0001545705454191193
step: 230, loss: 3.172348078805953e-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=0.985539488320356, f1=0.9777777777777777,
best_f1=0.9798206278026906
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                  : True
opt_level               : 02
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:14, 176.19it/s]
load_f1 = 0.9865771812080537
real_f1 = 0.9865771812080537
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, 226.56it/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: 0.6387677788734436  
step: 10, loss: 0.5025101900100708  
step: 20, loss: 0.4753020703792572  
step: 30, loss: 0.06845822185277939  
step: 40, loss: 0.17754705250263214  
step: 50, loss: 0.14768165349960327  
step: 60, loss: 0.05613873526453972  
step: 70, loss: 0.07321454584598541  
step: 80, loss: 0.023019013926386833  
step: 90, loss: 0.07988788932561874  
step: 100, loss: 0.00785745121538639  
step: 110, loss: 0.05053456127643585  
step: 120, loss: 0.1279267966747284  
step: 130, loss: 0.13486970961093903  
step: 140, loss: 0.09465517103672028  
step: 150, loss: 0.010164004750549793  
step: 160, loss: 0.013334548100829124  
step: 170, loss: 0.17911125719547272  
step: 180, loss: 0.04641527682542801  
step: 190, loss: 0.014130465686321259  
step: 200, loss: 0.14489972591400146  
step: 210, loss: 0.06461209058761597  
step: 220, loss: 0.17794233560562134  
step: 230, loss: 0.15901710093021393  
step: 240, loss: 0.053756825625896454  
step: 250, loss: 0.008495895192027092  
step: 260, loss: 0.028806094080209732  
step: 270, loss: 0.03672540932893753  
step: 280, loss: 0.04545047506690025  
step: 290, loss: 0.0949111357331276  
step: 300, loss: 0.025796623900532722  
step: 310, loss: 0.09689218550920486  
step: 320, loss: 0.12344147264957428  
step: 330, loss: 0.036980751901865005  
step: 340, loss: 0.02344975434243679  
step: 350, loss: 0.03151291608810425  
step: 360, loss: 0.006338948849588633  
step: 370, loss: 0.15442588925361633  
step: 380, loss: 0.012377185747027397  
step: 390, loss: 0.10988647490739822  
step: 400, loss: 0.17797406017780304  
step: 410, loss: 0.054396286606788635  
step: 420, loss: 0.024973131716251373  
step: 430, loss: 0.15297643840312958  
step: 440, loss: 0.014759457670152187  
step: 450, loss: 0.009946859441697598
```

```

step: 460, loss: 0.011162728071212769
step: 470, loss: 0.0970219150185585
step: 480, loss: 0.04013727232813835
step: 490, loss: 0.07227999716997147
step: 500, loss: 0.04903535172343254
step: 510, loss: 0.030463621020317078
step: 520, loss: 0.03992302343249321
step: 530, loss: 0.004000699147582054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9492787342950209, f1=0.9474662947466295,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2167222797870636
step: 10, loss: 0.03493497520685196
step: 20, loss: 0.029077734798192978
step: 30, loss: 0.09673546254634857
step: 40, loss: 0.12413590401411057
step: 50, loss: 0.13853804767131805
step: 60, loss: 0.004471127409487963
step: 70, loss: 0.01599668338894844
step: 80, loss: 0.06671983748674393
step: 90, loss: 0.008147332817316055
step: 100, loss: 0.005292470566928387
step: 110, loss: 0.004664616193622351
step: 120, loss: 0.028218021616339684
step: 130, loss: 0.04000385105609894
step: 140, loss: 0.012988969683647156
step: 150, loss: 0.023271022364497185
step: 160, loss: 0.0063752057030797005
step: 170, loss: 0.007854895666241646
step: 180, loss: 0.030024291947484016
step: 190, loss: 0.05821584165096283
step: 200, loss: 0.003280987963080406
step: 210, loss: 0.029087327420711517
step: 220, loss: 0.04416811838746071
step: 230, loss: 0.00427964236587286
step: 240, loss: 0.010217184200882912
step: 250, loss: 0.025660574436187744
step: 260, loss: 0.004084443207830191
step: 270, loss: 0.06290089339017868

```

```

step: 280, loss: 0.003345056436955929
step: 290, loss: 0.010830102488398552
step: 300, loss: 0.20243261754512787
step: 310, loss: 0.014951694756746292
step: 320, loss: 0.04421606659889221
step: 330, loss: 0.021444974467158318
step: 340, loss: 0.010184003971517086
step: 350, loss: 0.0008118590922094882
step: 360, loss: 0.09966595470905304
step: 370, loss: 0.0748746395111084
step: 380, loss: 0.031174510717391968
step: 390, loss: 0.023661447688937187
step: 400, loss: 0.04956592246890068
step: 410, loss: 0.06688886880874634
step: 420, loss: 0.0960191860795021
step: 430, loss: 0.01347440667450428
step: 440, loss: 0.2689056694507599
step: 450, loss: 0.021694457158446312
step: 460, loss: 0.019738933071494102
step: 470, loss: 0.09080386161804199
step: 480, loss: 0.2576000392436981
step: 490, loss: 0.013014709576964378
step: 500, loss: 0.24019190669059753
step: 510, loss: 0.010702740401029587
step: 520, loss: 0.10171540826559067
step: 530, loss: 0.04783646762371063
/usr/local/lib/python3.7/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=0.9483480688692415, f1=0.9474654377880185,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021160392090678215
step: 10, loss: 0.01634065806865692
step: 20, loss: 0.1545279175043106
step: 30, loss: 0.05721277371048927
step: 40, loss: 0.006296271458268166
step: 50, loss: 0.0038898452185094357
step: 60, loss: 0.004380091093480587
step: 70, loss: 0.003372886450961232
step: 80, loss: 0.0008028054144233465
step: 90, loss: 0.004811418242752552
step: 100, loss: 0.062049634754657745
step: 110, loss: 0.0017806708347052336
step: 120, loss: 0.0028686898294836283
step: 130, loss: 0.001610276522114873

```

```

step: 140, loss: 0.04627327620983124
step: 150, loss: 0.04592720419168472
step: 160, loss: 0.03662142902612686
step: 170, loss: 0.09605156630277634
step: 180, loss: 0.019061923027038574
step: 190, loss: 0.01728932186961174
step: 200, loss: 0.03639986738562584
step: 210, loss: 0.02884705550968647
step: 220, loss: 0.04547777771949768
step: 230, loss: 0.05858568102121353
step: 240, loss: 0.0008148823981173337
step: 250, loss: 0.01468848716467619
step: 260, loss: 0.004753207787871361
step: 270, loss: 0.0034449517261236906
step: 280, loss: 0.12410295009613037
step: 290, loss: 0.007795248180627823
step: 300, loss: 0.04085107147693634
step: 310, loss: 0.0018232968868687749
step: 320, loss: 0.022737614810466766
step: 330, loss: 0.0034137736074626446
step: 340, loss: 0.0027604938950389624
step: 350, loss: 0.0023485496640205383
step: 360, loss: 0.03736275061964989
step: 370, loss: 0.0043223244138062
step: 380, loss: 0.004147769883275032
step: 390, loss: 0.0021010416094213724
step: 400, loss: 0.029790310189127922
step: 410, loss: 0.045555680990219116
step: 420, loss: 0.1482631415128708
step: 430, loss: 0.009787577204406261
step: 440, loss: 0.0585797093808651
step: 450, loss: 0.06721073389053345
step: 460, loss: 0.10133273154497147
step: 470, loss: 0.15230031311511993
step: 480, loss: 0.00489908317103982
step: 490, loss: 0.004537977743893862
step: 500, loss: 0.0008865525014698505
step: 510, loss: 0.031812045723199844
step: 520, loss: 0.07793820649385452
step: 530, loss: 0.06738732755184174
/usr/local/lib/python3.7/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=0.9467838963442852, f1=0.9488243430152145,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.006482899654656649
step: 10, loss: 0.002634025178849697
step: 20, loss: 0.0008638736326247454
step: 30, loss: 0.0005463557317852974
step: 40, loss: 0.003839702345430851
step: 50, loss: 0.01025228388607502
step: 60, loss: 0.0011787363328039646
step: 70, loss: 0.0029418994672596455
step: 80, loss: 0.002289276337251067
step: 90, loss: 0.027485864236950874
step: 100, loss: 0.0014402546221390367
step: 110, loss: 0.0073998733423650265
step: 120, loss: 0.0002563111193012446
step: 130, loss: 0.0008078815299086273
step: 140, loss: 0.00046404197928495705
step: 150, loss: 0.031312402337789536
step: 160, loss: 0.017774058505892754
step: 170, loss: 0.0013011230621486902
step: 180, loss: 0.0005999921122565866
step: 190, loss: 0.026690369471907616
step: 200, loss: 0.0013405104400590062
step: 210, loss: 0.07225795835256577
step: 220, loss: 0.005923341028392315
step: 230, loss: 0.12322565168142319
step: 240, loss: 0.001015370711684227
step: 250, loss: 0.0018117667641490698
step: 260, loss: 0.0018497892888262868
step: 270, loss: 0.004665741231292486
step: 280, loss: 0.009344219230115414
step: 290, loss: 0.03738775849342346
step: 300, loss: 0.0001091750746127218
step: 310, loss: 0.0014249109663069248
step: 320, loss: 0.0005544201121665537
step: 330, loss: 0.05599430575966835
step: 340, loss: 0.013604787178337574
step: 350, loss: 0.0013695331290364265
step: 360, loss: 0.04648977890610695
step: 370, loss: 0.002553569385781884
step: 380, loss: 0.11070238798856735
step: 390, loss: 0.0007604664424434304
step: 400, loss: 0.0008828417048789561
step: 410, loss: 0.003100191941484809
step: 420, loss: 0.005321017000824213
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 0.015625884756445885
step: 440, loss: 0.0007361255120486021
step: 450, loss: 0.11775597929954529
step: 460, loss: 0.00022542184160556644

```

step: 470, loss: 0.0012497958960011601
step: 480, loss: 0.07511337846517563
step: 490, loss: 0.0029674989636987448
step: 500, loss: 0.003673827275633812
step: 510, loss: 0.0029644279275089502
step: 520, loss: 0.0697089433670044
step: 530, loss: 0.03227962926030159
/usr/local/lib/python3.7/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=0.9463955637707948, f1=0.9422632794457274,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00489429896697402
step: 10, loss: 0.03397085517644882
step: 20, loss: 0.0016219126991927624
step: 30, loss: 0.00024457642575725913
step: 40, loss: 0.008099466562271118
step: 50, loss: 0.002756169531494379
step: 60, loss: 0.01652839593589306
step: 70, loss: 0.0034421521704643965
step: 80, loss: 0.00018715033365879208
step: 90, loss: 0.0019053551368415356
step: 100, loss: 0.0005493300268426538
step: 110, loss: 0.0003469662624411285
step: 120, loss: 0.000400111370254308
step: 130, loss: 0.0012419546255841851
step: 140, loss: 0.0017372403526678681
step: 150, loss: 0.00218955404125154
step: 160, loss: 0.0006439726566895843
step: 170, loss: 0.02382407896220684
step: 180, loss: 0.00080135278403759
step: 190, loss: 0.0017378153279423714
step: 200, loss: 0.0005411286256276071
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 0.002592650009319186
step: 220, loss: 0.0022387036588042974
step: 230, loss: 0.008358610793948174
step: 240, loss: 0.005030290223658085
step: 250, loss: 0.0035948820877820253
step: 260, loss: 0.0024114344269037247
step: 270, loss: 0.0003726449212990701
step: 280, loss: 0.0015643443912267685
step: 290, loss: 0.12640510499477386
step: 300, loss: 0.004301451146602631
step: 310, loss: 0.0011006247950717807

```



```

step: 320, loss: 0.11295929551124573
step: 330, loss: 0.03432651609182358
step: 340, loss: 0.024321096017956734
step: 350, loss: 0.0021356341894716024
step: 360, loss: 0.0012548784725368023
step: 370, loss: 0.0019111859146505594
step: 380, loss: 0.00013617869990412146
step: 390, loss: 9.694363689050078e-05
step: 400, loss: 0.004576413426548243
step: 410, loss: 0.016570940613746643
step: 420, loss: 0.00032305155764333904
step: 430, loss: 0.000716346490662545
step: 440, loss: 0.004591597709804773
step: 450, loss: 0.000453750544693321
step: 460, loss: 0.0001925716205732897
step: 470, loss: 0.007640228606760502
step: 480, loss: 0.0008262646733783185
step: 490, loss: 0.00016371585661545396
step: 500, loss: 0.001499759848229587
step: 510, loss: 0.004425754304975271
step: 520, loss: 0.0005035821814090014
step: 530, loss: 0.0017597842961549759
/usr/local/lib/python3.7/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=0.9430379746835444, f1=0.9358456707043517,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08837846666574478
step: 10, loss: 0.10019215941429138
step: 20, loss: 0.06655971705913544
step: 30, loss: 0.022202862426638603
step: 40, loss: 0.004678352270275354
step: 50, loss: 0.025199487805366516
step: 60, loss: 0.00010561761155258864
step: 70, loss: 0.0002877590886782855
step: 80, loss: 0.0007389994570985436
step: 90, loss: 0.00022877789160702378
step: 100, loss: 0.007707113400101662
step: 110, loss: 0.0038770907558500767
step: 120, loss: 0.00047001821803860366
step: 130, loss: 0.0365850031375885
step: 140, loss: 0.005867107771337032
step: 150, loss: 0.0003701520327012986
step: 160, loss: 0.00017360530910082161
step: 170, loss: 0.000122549245133996

```

```

step: 180, loss: 0.001834760420024395
step: 190, loss: 0.00010607259901007637
step: 200, loss: 0.0001625090662855655
step: 210, loss: 0.0011992547661066055
step: 220, loss: 0.00032247090712189674
step: 230, loss: 0.0008164052851498127
step: 240, loss: 0.04702196642756462
step: 250, loss: 0.0014406415866687894
step: 260, loss: 0.002706534694880247
step: 270, loss: 0.08496365696191788
step: 280, loss: 0.003322375938296318
step: 290, loss: 0.00031311577185988426
step: 300, loss: 0.008829692378640175
step: 310, loss: 0.001605289988219738
step: 320, loss: 0.0014057205989956856
step: 330, loss: 0.0010927296243607998
step: 340, loss: 0.034255992621183395
step: 350, loss: 0.014644360169768333
step: 360, loss: 0.016831140965223312
step: 370, loss: 0.0034754585940390825
step: 380, loss: 0.00013221142580732703
step: 390, loss: 0.0996895581483841
step: 400, loss: 0.0012920988956466317
step: 410, loss: 0.01595139503479004
step: 420, loss: 0.000824999064207077
step: 430, loss: 0.0003553265705704689
step: 440, loss: 8.700860780663788e-05
step: 450, loss: 0.014465468004345894
step: 460, loss: 0.002042242791503668
step: 470, loss: 0.0009087977814488113
step: 480, loss: 0.05145489424467087
step: 490, loss: 0.002100268378853798
step: 500, loss: 0.006350615527480841
step: 510, loss: 0.0029474811162799597
step: 520, loss: 0.020812565460801125
step: 530, loss: 0.02422475256025791
/usr/local/lib/python3.7/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=0.9476609541454377, f1=0.9470290188853063,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001073458231985569
step: 10, loss: 0.0003685908450279385
step: 20, loss: 0.0005518245161511004
step: 30, loss: 0.004246389959007502

```

step: 40, loss: 0.023275982588529587
step: 50, loss: 0.002850133227184415
step: 60, loss: 0.0016026583034545183
step: 70, loss: 0.00011913864727830514
step: 80, loss: 0.0006608051480725408
step: 90, loss: 0.0005392778548412025
step: 100, loss: 4.485460522118956e-05
step: 110, loss: 0.0005546597531065345
step: 120, loss: 0.00010582974209683016
step: 130, loss: 0.000565864727832377
step: 140, loss: 0.00011073299538111314
step: 150, loss: 0.0057379091158509254
step: 160, loss: 9.201789362123236e-05
step: 170, loss: 8.77127458807081e-05
step: 180, loss: 0.0002465460856910795
step: 190, loss: 0.00016648830205667764
step: 200, loss: 7.539763464592397e-05
step: 210, loss: 0.00029769266257062554
step: 220, loss: 6.976172153372318e-05
step: 230, loss: 0.0002820160589180887
step: 240, loss: 0.0001574061025166884
step: 250, loss: 0.0005015206406824291
step: 260, loss: 4.806180368177593e-05
step: 270, loss: 0.0005008849548175931
step: 280, loss: 5.9666112065315247e-05
step: 290, loss: 5.8337820519227535e-05
step: 300, loss: 0.00016744632739573717
step: 310, loss: 8.715814328752458e-05
step: 320, loss: 9.522488835500553e-05
step: 330, loss: 0.004014521837234497
step: 340, loss: 0.019108619540929794
step: 350, loss: 0.0017319538164883852
step: 360, loss: 0.0005833720788359642
step: 370, loss: 0.00033032408100552857
step: 380, loss: 0.01176801510155201
step: 390, loss: 0.001210540416650474
step: 400, loss: 0.0002175225963583216
step: 410, loss: 0.0005259652971290052
step: 420, loss: 0.0006692685419693589
step: 430, loss: 0.0001812246482586488
step: 440, loss: 0.0018805747386068106
step: 450, loss: 0.0007327857310883701
step: 460, loss: 0.0036982304882258177
step: 470, loss: 0.00018385903968010098
step: 480, loss: 0.1451871693134308
step: 490, loss: 0.001486724242568016
step: 500, loss: 0.004453470930457115
step: 510, loss: 0.0010841782204806805

```

step: 520, loss: 0.004061136394739151
step: 530, loss: 0.005446469411253929
/usr/local/lib/python3.7/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=0.9478138222849083, f1=0.9439775910364145,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015957903815433383
step: 10, loss: 0.045496758073568344
step: 20, loss: 0.00039114319952204823
step: 30, loss: 0.0002695831935852766
step: 40, loss: 6.72275127726607e-05
step: 50, loss: 0.00875181332230568
step: 60, loss: 0.00046881806338205934
step: 70, loss: 9.802151180338115e-05
step: 80, loss: 0.00040502040064893663
step: 90, loss: 0.000267350667854771
step: 100, loss: 7.062803342705593e-05
step: 110, loss: 0.0008850517915561795
step: 120, loss: 0.00011684893252095208
step: 130, loss: 8.370012801606208e-05
step: 140, loss: 0.010706465691328049
step: 150, loss: 4.2860621761064976e-05
step: 160, loss: 0.0012235920876264572
step: 170, loss: 0.0005289234104566276
step: 180, loss: 4.814802741748281e-05
step: 190, loss: 0.0011896949727088213
step: 200, loss: 0.0011071910848841071
step: 210, loss: 3.833204027614556e-05
step: 220, loss: 0.002793773775920272
step: 230, loss: 7.395105058094487e-05
step: 240, loss: 0.0009398465044796467
step: 250, loss: 0.006105619482696056
step: 260, loss: 6.511791434604675e-05
step: 270, loss: 0.0047095469199121
step: 280, loss: 0.0029939296655356884
step: 290, loss: 0.0003076343273278326
step: 300, loss: 0.00019643804989755154
step: 310, loss: 0.003913033287972212
step: 320, loss: 0.023815156891942024
step: 330, loss: 0.04771249741315842
step: 340, loss: 0.0014148758491501212
step: 350, loss: 3.6826928408117965e-05
step: 360, loss: 0.0008593285456299782
step: 370, loss: 0.00031832512468099594

```

```

step: 380, loss: 0.037952858954668045
step: 390, loss: 0.04791926220059395
step: 400, loss: 0.0001864410296548158
step: 410, loss: 8.83187385625206e-05
step: 420, loss: 0.0009057121351361275
step: 430, loss: 0.04057302325963974
step: 440, loss: 0.0011796151520684361
step: 450, loss: 0.00010940522770397365
step: 460, loss: 6.558478344231844e-05
step: 470, loss: 0.0011827758280560374
step: 480, loss: 4.754257679451257e-05
step: 490, loss: 0.008193189278244972
step: 500, loss: 0.000841658329591155
step: 510, loss: 0.0002299602929269895
step: 520, loss: 0.001852164277806878
step: 530, loss: 0.00023707230866421014
/usr/local/lib/python3.7/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=0.9490566037735849, f1=0.9472693032015067,
best_f1=0.9474662947466295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017353685107082129
step: 10, loss: 0.0010948985582217574
step: 20, loss: 0.00044773094123229384
step: 30, loss: 0.0032106193248182535
step: 40, loss: 2.8400820156093687e-05
step: 50, loss: 2.6552623239695095e-05
step: 60, loss: 5.530704220291227e-05
step: 70, loss: 9.654599125497043e-05
step: 80, loss: 0.0026682072784751654
step: 90, loss: 5.666774814017117e-05
step: 100, loss: 8.472029003314674e-05
step: 110, loss: 0.000224019298912026
step: 120, loss: 0.00020384276285767555
step: 130, loss: 0.0003558204334694892
step: 140, loss: 0.029342474415898323
step: 150, loss: 0.0001129216470872052
step: 160, loss: 0.00015696813352406025
step: 170, loss: 0.004263211973011494
step: 180, loss: 7.359572919085622e-05
step: 190, loss: 0.0009453041129745543
step: 200, loss: 0.0022583846002817154
step: 210, loss: 0.00015844005974940956
step: 220, loss: 0.0013959029456600547
step: 230, loss: 8.067159069469199e-05

```

```

step: 240, loss: 6.889036012580618e-05
step: 250, loss: 0.0016843373887240887
step: 260, loss: 0.0037758364342153072
step: 270, loss: 2.9514361813198775e-05
step: 280, loss: 0.0016571675660088658
step: 290, loss: 0.02839563600718975
step: 300, loss: 3.446865230216645e-05
step: 310, loss: 0.0062741669826209545
step: 320, loss: 0.0001716283441055566
step: 330, loss: 0.0004982731770724058
step: 340, loss: 3.095576539635658e-05
step: 350, loss: 6.406063766917214e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 360, loss: 2.8158170607639477e-05
step: 370, loss: 0.001121297711506486
step: 380, loss: 3.6838097003055736e-05
step: 390, loss: 2.870616845029872e-05
step: 400, loss: 0.0008142748847603798
step: 410, loss: 2.6444920877111144e-05
step: 420, loss: 3.078485678997822e-05
step: 430, loss: 4.5152562961447984e-05
step: 440, loss: 0.0006641875370405614
step: 450, loss: 3.123146962025203e-05
step: 460, loss: 7.003530481597409e-05
step: 470, loss: 2.1952553652226925e-05
step: 480, loss: 5.394459003582597e-05
step: 490, loss: 0.017100034281611443
step: 500, loss: 8.834942855173722e-05
step: 510, loss: 0.0019792867824435234
step: 520, loss: 0.0005468474701046944
step: 530, loss: 0.00013645217404700816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9559512652296158, f1=0.9496268656716418,
best_f1=0.9496268656716418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.414086677366868e-05
step: 10, loss: 4.9529138777870685e-05
step: 20, loss: 1.9575976693886332e-05
step: 30, loss: 6.606060196645558e-05
step: 40, loss: 7.105860277079046e-05

```

step: 50, loss: 0.00014112335338722914
step: 60, loss: 7.492233999073505e-05
step: 70, loss: 0.0014593950472772121
step: 80, loss: 3.9553091482957825e-05
step: 90, loss: 0.00016179689555428922
step: 100, loss: 0.0002555248502176255
step: 110, loss: 3.030392326763831e-05
step: 120, loss: 0.0002668368979357183
step: 130, loss: 0.00011852899478981271
step: 140, loss: 0.006390288472175598
step: 150, loss: 0.00035141006810590625
step: 160, loss: 5.096658060210757e-05
step: 170, loss: 6.239629874471575e-05
step: 180, loss: 0.0003454501857049763
step: 190, loss: 0.002549537690356374
step: 200, loss: 0.002088323701173067
step: 210, loss: 0.00042444971040822566
step: 220, loss: 0.0014483437407761812
step: 230, loss: 0.0002921014674939215
step: 240, loss: 6.125690561020747e-05
step: 250, loss: 0.00012013975356239825
step: 260, loss: 0.002202199539169669
step: 270, loss: 0.00011956765956711024
step: 280, loss: 3.656945045804605e-05
step: 290, loss: 0.00012221607903484255
step: 300, loss: 0.0001026520985760726
step: 310, loss: 4.603144043358043e-05
step: 320, loss: 0.003292744979262352
step: 330, loss: 9.282852261094376e-05
step: 340, loss: 0.0005969193298369646
step: 350, loss: 9.987833618652076e-05
step: 360, loss: 0.0006963419145904481
step: 370, loss: 0.0023020156659185886
step: 380, loss: 0.0064339544624090195
step: 390, loss: 0.000615412020124495
step: 400, loss: 0.001244189334101975
step: 410, loss: 0.000602273503318429
step: 420, loss: 0.0001540036464575678
step: 430, loss: 3.377089524292387e-05
step: 440, loss: 0.0014783896040171385
step: 450, loss: 0.00014815517351962626
step: 460, loss: 0.00032791864941827953
step: 470, loss: 0.000674057868309319
step: 480, loss: 5.575093746301718e-05
step: 490, loss: 0.05885028839111328
step: 500, loss: 0.0007312042871490121
step: 510, loss: 9.908556239679456e-05
step: 520, loss: 0.00011958803224842995

```

step: 530, loss: 5.02750335726887e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.951048951048951, f1=0.9447795823665893,
best_f1=0.9496268656716418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024808195303194225
step: 10, loss: 0.13804295659065247
step: 20, loss: 0.003053725929930806
step: 30, loss: 0.00083550886483863
step: 40, loss: 0.0036380679812282324
step: 50, loss: 0.0004245475574862212
step: 60, loss: 0.013882513158023357
step: 70, loss: 0.00017950832261703908
step: 80, loss: 3.136930172331631e-05
step: 90, loss: 9.870580106507987e-05
step: 100, loss: 0.0004952654126100242
step: 110, loss: 6.294492050074041e-05
step: 120, loss: 3.196157922502607e-05
step: 130, loss: 3.1208470318233594e-05
step: 140, loss: 7.705426105530933e-05
step: 150, loss: 3.5083809052594006e-05
step: 160, loss: 4.0178052586270496e-05
step: 170, loss: 3.97683288611006e-05
step: 180, loss: 0.00010498146730242297
step: 190, loss: 0.00040315999649465084
step: 200, loss: 0.0007884094957262278
step: 210, loss: 0.0583450049161911
step: 220, loss: 0.0002495681692380458
step: 230, loss: 0.00010406469664303586
step: 240, loss: 0.0001756417186697945
step: 250, loss: 1.7102554920711555e-05
step: 260, loss: 0.00020101116388104856
step: 270, loss: 0.0006587597890757024
step: 280, loss: 5.299084659782238e-05
step: 290, loss: 3.9835220377426594e-05
step: 300, loss: 2.8407781428541057e-05
step: 310, loss: 5.002061152481474e-05
step: 320, loss: 0.0007481356151401997
step: 330, loss: 0.0006474233232438564
step: 340, loss: 3.0262459404184483e-05
step: 350, loss: 0.05723197013139725
step: 360, loss: 3.78815529984422e-05
step: 370, loss: 0.0002325720270164311
step: 380, loss: 3.567372550605796e-05

```



```

step: 390, loss: 0.00017603952437639236
step: 400, loss: 0.0001616164081497118
step: 410, loss: 2.7167898224433884e-05
step: 420, loss: 8.162696030922234e-05
step: 430, loss: 4.2712767026387155e-05
step: 440, loss: 6.395963282557204e-05
step: 450, loss: 0.00013322119775693864
step: 460, loss: 0.022202247753739357
step: 470, loss: 2.6795156372827478e-05
step: 480, loss: 5.5897238780744374e-05
step: 490, loss: 0.00012912503734696656
step: 500, loss: 5.5224427342182025e-05
step: 510, loss: 4.928979979013093e-05
step: 520, loss: 0.034346383064985275
step: 530, loss: 2.149449755961541e-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=0.9554317548746518, f1=0.948291782086796,
best_f1=0.9496268656716418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.816611777234357e-05
step: 10, loss: 1.999310916289687e-05
step: 20, loss: 8.875412459019572e-05
step: 30, loss: 0.0016952778678387403
step: 40, loss: 5.018917363486253e-05
step: 50, loss: 0.03611735254526138
step: 60, loss: 0.00014218140859156847
step: 70, loss: 2.7607326046563685e-05
step: 80, loss: 3.8828067772556096e-05
step: 90, loss: 0.00016961607616394758
step: 100, loss: 2.6537700250628404e-05
step: 110, loss: 2.3032935132505372e-05
step: 120, loss: 1.715095822873991e-05
step: 130, loss: 0.0001748222130117938
step: 140, loss: 1.8674607417779043e-05
step: 150, loss: 2.134161513822619e-05
step: 160, loss: 1.5258605344570242e-05
step: 170, loss: 2.305120688106399e-05
step: 180, loss: 0.00041505167610011995
step: 190, loss: 3.145676964777522e-05
step: 200, loss: 4.2221363401040435e-05
step: 210, loss: 0.00014402292435988784
step: 220, loss: 4.6940749598434195e-05
step: 230, loss: 1.584341771376785e-05
step: 240, loss: 3.9737926272209734e-05

```

```

step: 250, loss: 1.928905476233922e-05
step: 260, loss: 3.246497726649977e-05
step: 270, loss: 2.5487892344244756e-05
step: 280, loss: 3.474598270258866e-05
step: 290, loss: 2.744203811744228e-05
step: 300, loss: 0.009993325918912888
step: 310, loss: 1.9278084437246434e-05
step: 320, loss: 2.2775717297918163e-05
step: 330, loss: 2.2455622456618585e-05
step: 340, loss: 2.1311976524884813e-05
step: 350, loss: 3.4486281947465613e-05
step: 360, loss: 4.3389816710259765e-05
step: 370, loss: 8.922544657252729e-05
step: 380, loss: 2.4109569494612515e-05
step: 390, loss: 0.000457615009509027
step: 400, loss: 1.9833145415759645e-05
step: 410, loss: 3.0424391297856346e-05
step: 420, loss: 0.003721675369888544
step: 430, loss: 0.014967411756515503
step: 440, loss: 0.0002659888123162091
step: 450, loss: 0.0428137443959713
step: 460, loss: 1.6901405615499243e-05
step: 470, loss: 0.04070838540792465
step: 480, loss: 0.0001117233041441068
step: 490, loss: 0.0001684598537394777
step: 500, loss: 7.309168722713366e-05
step: 510, loss: 0.18677978217601776
step: 520, loss: 2.2846546926302835e-05
step: 530, loss: 0.06988989561796188
/usr/local/lib/python3.7/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=0.9522918615528532, f1=0.9413953488372093,
best_f1=0.9496268656716418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001505822583567351
step: 10, loss: 1.8875689420383424e-05
step: 20, loss: 3.350921178935096e-05
step: 30, loss: 0.0004459074407350272
step: 40, loss: 2.0749428585986607e-05
step: 50, loss: 5.150417564436793e-05
step: 60, loss: 0.0015025873435661197
step: 70, loss: 1.470725783292437e-05
step: 80, loss: 2.1799247406306677e-05
step: 90, loss: 0.010055270045995712
step: 100, loss: 0.0004590458411257714

```

```
step: 110, loss: 2.3404520106851123e-05
step: 120, loss: 0.00034783067530952394
step: 130, loss: 3.631512663559988e-05
step: 140, loss: 0.00015768605226185173
step: 150, loss: 0.000370911555364728
step: 160, loss: 0.0005003145197406411
step: 170, loss: 0.0004960006917826831
step: 180, loss: 4.062161315232515e-05
step: 190, loss: 5.355905886972323e-05
step: 200, loss: 2.1013869627495296e-05
step: 210, loss: 2.5553395971655846e-05
step: 220, loss: 1.5202674148895312e-05
step: 230, loss: 1.352636536466889e-05
step: 240, loss: 0.00020297399896662682
step: 250, loss: 1.7199365174747072e-05
step: 260, loss: 3.05522044072859e-05
step: 270, loss: 1.6141366359079257e-05
step: 280, loss: 2.1978383301757276e-05
step: 290, loss: 2.903178938140627e-05
step: 300, loss: 2.372955896134954e-05
step: 310, loss: 0.0010230778716504574
step: 320, loss: 0.008684728294610977
step: 330, loss: 0.0010677219834178686
step: 340, loss: 2.1836674932274036e-05
step: 350, loss: 1.9736149624804966e-05
step: 360, loss: 1.8387718228041194e-05
step: 370, loss: 5.045188299845904e-05
step: 380, loss: 0.00013351008237805218
step: 390, loss: 0.00019808707293123007
step: 400, loss: 4.389758032630198e-05
step: 410, loss: 0.0001971954043256119
step: 420, loss: 1.688637894403655e-05
step: 430, loss: 0.0012029237113893032
step: 440, loss: 5.5584048823220655e-05
step: 450, loss: 3.5002176446141675e-05
step: 460, loss: 0.0003679264336824417
step: 470, loss: 1.767204412317369e-05
step: 480, loss: 1.9683962818817236e-05
step: 490, loss: 2.721163400565274e-05
step: 500, loss: 1.4360701243276708e-05
step: 510, loss: 2.4440236302325502e-05
step: 520, loss: 0.0006256606429815292
step: 530, loss: 1.2758993761963211e-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=0.9536299765807962, f1=0.9455560725919032,
best_f1=0.9496268656716418
```

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

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

step: 0, loss: 2.079410114674829e-05
step: 10, loss: 1.180161507363664e-05
step: 20, loss: 1.3645617400470655e-05
step: 30, loss: 5.6738706916803494e-05
step: 40, loss: 3.9629660022910684e-05
step: 50, loss: 0.00010416915029054508
step: 60, loss: 7.002293568802997e-05
step: 70, loss: 4.6795044909231365e-05
step: 80, loss: 1.751976196828764e-05
step: 90, loss: 3.1459261663258076e-05
step: 100, loss: 6.713892798870802e-05
step: 110, loss: 5.277035234030336e-05
step: 120, loss: 6.03845837758854e-05
step: 130, loss: 0.0159321092069149
step: 140, loss: 3.827216278295964e-05
step: 150, loss: 1.287072245759191e-05
step: 160, loss: 1.857764982560184e-05
step: 170, loss: 0.0006508665392175317
step: 180, loss: 0.0006525029893964529
step: 190, loss: 2.5516905225231312e-05
step: 200, loss: 2.6295949282939546e-05
step: 210, loss: 1.4174590432958212e-05
step: 220, loss: 1.8179032849729992e-05
step: 230, loss: 0.0007508869166485965
step: 240, loss: 4.572247780743055e-05
step: 250, loss: 2.0551808120217174e-05
step: 260, loss: 3.3110765798483044e-05
step: 270, loss: 2.3225769837154076e-05
step: 280, loss: 1.3049577319179662e-05
step: 290, loss: 7.024969818303362e-05
step: 300, loss: 2.8832248062826693e-05
step: 310, loss: 3.645868491730653e-05
step: 320, loss: 5.801334191346541e-05
step: 330, loss: 5.796054028905928e-05
step: 340, loss: 1.858516588981729e-05
step: 350, loss: 0.0012011905200779438
step: 360, loss: 0.000173769862158224
step: 370, loss: 1.7247781215701252e-05
step: 380, loss: 1.6566125850658864e-05
step: 390, loss: 0.025510597974061966
step: 400, loss: 4.141787576372735e-05
step: 410, loss: 2.011574360949453e-05
step: 420, loss: 2.2603313482250087e-05
step: 430, loss: 2.201210372732021e-05
step: 440, loss: 4.9857524572871625e-05

```

step: 450, loss: 3.126112642348744e-05
step: 460, loss: 3.0015553420525976e-05
step: 470, loss: 3.9424572605639696e-05
step: 480, loss: 2.2730851924279705e-05
step: 490, loss: 2.3929525923449546e-05
step: 500, loss: 1.6457894162158482e-05
step: 510, loss: 8.577304834034294e-05
step: 520, loss: 2.2991374862613156e-05
step: 530, loss: 2.2041771444492042e-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=0.9545667447306792, f1=0.9464368886818817,
best_f1=0.9496268656716418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.814540519262664e-05
step: 10, loss: 2.1251511498121545e-05
step: 20, loss: 1.9523482478689402e-05
step: 30, loss: 8.924463327275589e-05
step: 40, loss: 0.02758924663066864
step: 50, loss: 2.09466179512674e-05
step: 60, loss: 1.1853718206111807e-05
step: 70, loss: 0.0018682939698919654
step: 80, loss: 2.4514902179362252e-05
step: 90, loss: 1.5225098650262225e-05
step: 100, loss: 6.406143074855208e-05
step: 110, loss: 2.0075183783774264e-05
step: 120, loss: 2.6362138669355772e-05
step: 130, loss: 0.05604758858680725
step: 140, loss: 1.233802322531119e-05
step: 150, loss: 2.550908538978547e-05
step: 160, loss: 1.4066522453504149e-05
step: 170, loss: 1.4230420674721245e-05
step: 180, loss: 1.0121535524376668e-05
step: 190, loss: 1.8789996829582378e-05
step: 200, loss: 1.5619798432453535e-05
step: 210, loss: 2.247007068945095e-05
step: 220, loss: 1.135829916165676e-05
step: 230, loss: 1.4394316167454235e-05
step: 240, loss: 4.755979898618534e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 250, loss: 1.5802401321707293e-05
step: 260, loss: 8.946289017330855e-05
step: 270, loss: 3.2011885195970535e-05
step: 280, loss: 6.335886428132653e-05
step: 290, loss: 1.1745723895728588e-05

```

```

step: 300, loss: 2.7416130251367576e-05
step: 310, loss: 0.0031937966123223305
step: 320, loss: 1.6167146895895712e-05
step: 330, loss: 9.288005821872503e-05
step: 340, loss: 1.5172825442277826e-05
step: 350, loss: 1.2766395229846239e-05
step: 360, loss: 6.350333569571376e-05
step: 370, loss: 1.4785413441131823e-05
step: 380, loss: 3.5338580346433446e-05
step: 390, loss: 0.0001840459299273789
step: 400, loss: 1.4170846043271013e-05
step: 410, loss: 3.4243243135279045e-05
step: 420, loss: 5.225673521636054e-05
step: 430, loss: 0.0006843549199402332
step: 440, loss: 4.355085184215568e-05
step: 450, loss: 1.2680727195402142e-05
step: 460, loss: 1.6256861272267997e-05
step: 470, loss: 3.080965325352736e-05
step: 480, loss: 3.184332308592275e-05
step: 490, loss: 1.2185308150947094e-05
step: 500, loss: 2.0760293409693986e-05
step: 510, loss: 7.007042586337775e-05
step: 520, loss: 1.359344332740875e-05
step: 530, loss: 8.543998410459608e-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=0.9531757070004635, f1=0.9444189251263206,
best_f1=0.9496268656716418
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 238.19it/s]
load_f1 = 0.9555035128805621
real_f1 = 0.9550140581068416
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 238.11it/s]

```

3 DK

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                                     | 827.9 MB 532 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: 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: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.6)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in

```

```

/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: 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: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: 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: 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: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: srsly<1.1.0,>=0.1.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: 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=6ee73eb936a64c5178761c884c2083182be39db1119f3227b97648e633c4df9
  Stored in directory: /tmp/pip-ephem-wheel-cache-67nq7t88/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/
```

```
Downloading: 100% 433/433 [00:00<00:00, 395kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 340kB/s]
Downloading: 100% 440M/440M [00:06<00:00, 71.6MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

```
Defaults for this optimization level are:
```

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

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

```
After processing overrides, optimization options are:
```

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

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

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

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

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

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

```
step: 0, loss: 0.5504838228225708
```

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

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

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

```
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=0.45, f1=0.46511627906976755, best_f1=0.46511627906976755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4863928556442261
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6000000000000001, f1=0.5000000000000001,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4317949414253235
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.4864864864864865,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3134617507457733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.6666666666666666, f1=0.5517241379310344,
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: 0.2786325216293335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7857142857142857, f1=0.7333333333333334,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2460520714521408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8666666666666666, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02347414381802082
/usr/local/lib/python3.7/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=0.7777777777777778, f1=0.8235294117647058,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011837170459330082
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8571428571428571,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017115281661972404
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.782608695652174,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006979288533329964
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8666666666666666,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032680686563253403
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.8666666666666666,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0042025926522910595
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.8888888888888889,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038364927750080824
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.8888888888888889,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002847085241228342
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.8888888888888889,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026417721528559923
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.8888888888888889,

```

```

best_f1=0.7741935483870968
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, 130355.76it/s]
/usr/local/lib/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 = 0.896551724137931
real_f1 = 0.896551724137931
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 200.98it/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: 0.5886775851249695
step: 10, loss: 0.6111385226249695
step: 20, loss: 0.25244206190109253
step: 30, loss: 0.12714381515979767
step: 40, loss: 0.24371325969696045
step: 50, loss: 0.017230695113539696
step: 60, loss: 0.014592322520911694
step: 70, loss: 0.0744144544005394
step: 80, loss: 0.09605907648801804
step: 90, loss: 0.16055649518966675
step: 100, loss: 0.005079512018710375
step: 110, loss: 0.14585080742835999
step: 120, loss: 0.0750761404633522
step: 130, loss: 0.07516144961118698
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.007406463846564293
step: 150, loss: 0.016188545152544975
step: 160, loss: 0.01406849268823862
step: 170, loss: 0.08398904651403427
step: 180, loss: 0.00233708624728024
step: 190, loss: 0.07653526216745377
step: 200, loss: 0.016065174713730812
step: 210, loss: 0.01040572114288807
step: 220, loss: 0.05177713930606842
step: 230, loss: 0.01493796892464161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876543209876544, f1=0.9898534385569334,
best_f1=0.9898534385569334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016872297273948789
step: 10, loss: 0.004858808126300573
step: 20, loss: 0.02105654962360859
step: 30, loss: 0.2029794156551361
step: 40, loss: 0.039980728179216385
step: 50, loss: 0.0011730385012924671
step: 60, loss: 0.0009520367020741105
step: 70, loss: 0.07626799494028091
step: 80, loss: 0.005536686163395643
step: 90, loss: 0.0053803869523108006
step: 100, loss: 0.003250165842473507
step: 110, loss: 0.08349426835775375
step: 120, loss: 0.0006576123996637762
step: 130, loss: 0.004513956606388092
step: 140, loss: 0.001612642896361649
step: 150, loss: 0.0024274957831948996
step: 160, loss: 0.0019877918530255556
step: 170, loss: 0.0005843131220899522
step: 180, loss: 0.0053733838722109795
step: 190, loss: 0.001913245301693678
step: 200, loss: 0.0010261785937473178
step: 210, loss: 0.0007740020519122481
step: 220, loss: 0.09962789714336395
step: 230, loss: 0.002544857095927
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.990990990990991,
best_f1=0.9898534385569334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015857269754633307
step: 10, loss: 0.002088935114443302
step: 20, loss: 0.000528931908775121
step: 30, loss: 0.03941668942570686
step: 40, loss: 0.12894411385059357
step: 50, loss: 0.004374152515083551
step: 60, loss: 0.002053221920505166
step: 70, loss: 0.0030433402862399817
step: 80, loss: 0.003256717463955283
step: 90, loss: 0.05791837349534035
step: 100, loss: 0.0008332219440490007

```

```

step: 110, loss: 0.0019465284422039986
step: 120, loss: 0.039498355239629745
step: 130, loss: 0.003735723439604044
step: 140, loss: 0.0046080294996500015
step: 150, loss: 0.08008404076099396
step: 160, loss: 0.01580694317817688
step: 170, loss: 0.01478307880461216
step: 180, loss: 0.017366571351885796
step: 190, loss: 0.004666720051318407
step: 200, loss: 0.0026987942401319742
step: 210, loss: 0.0007557274657301605
step: 220, loss: 0.0006625294918194413
step: 230, loss: 0.0003977056476287544
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9898762654668166,
best_f1=0.9898534385569334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031784988823346794
step: 10, loss: 0.0019391621463000774
step: 20, loss: 0.0014826025580987334
step: 30, loss: 0.0013349715154618025
step: 40, loss: 0.0013558408245444298
step: 50, loss: 0.0005534880328923464
step: 60, loss: 0.0018984676571562886
step: 70, loss: 0.0003013512759935111
step: 80, loss: 0.006464628968387842
step: 90, loss: 0.007839656434953213
step: 100, loss: 0.0002227185614174232
step: 110, loss: 0.0004167981678619981
step: 120, loss: 0.047569308429956436
step: 130, loss: 0.002032620133832097
step: 140, loss: 0.0012550422688946128
step: 150, loss: 0.10111510008573532
step: 160, loss: 0.06619405746459961
step: 170, loss: 0.02118103764951229
step: 180, loss: 0.0011024409905076027
step: 190, loss: 0.011214549653232098
step: 200, loss: 0.004259572830051184
step: 210, loss: 0.09540606290102005
step: 220, loss: 0.0004612256307154894
step: 230, loss: 0.014586854726076126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.988814317673378, f1=0.9876543209876544,
best_f1=0.9876543209876544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007695852778851986
step: 10, loss: 0.0003644484968390316
step: 20, loss: 0.00086480617756024
step: 30, loss: 0.002410331042483449
step: 40, loss: 0.00035821174969896674
step: 50, loss: 0.00023437326308339834
step: 60, loss: 0.0007665111334063113
step: 70, loss: 0.0005818227073177695
step: 80, loss: 0.00026597658870741725
step: 90, loss: 0.0008267919765785336
step: 100, loss: 0.0004382259212434292
step: 110, loss: 0.00021709446446038783
step: 120, loss: 9.89311738521792e-05
step: 130, loss: 0.00043701587128452957
step: 140, loss: 0.00020543605205602944
step: 150, loss: 0.0004254111845511943
step: 160, loss: 0.03184715285897255
step: 170, loss: 0.010371428914368153
step: 180, loss: 0.0014461437240242958
step: 190, loss: 0.014756842516362667
step: 200, loss: 0.0005666989600285888
step: 210, loss: 0.00043180419015698135
step: 220, loss: 0.004334527999162674
step: 230, loss: 0.0003502287727314979
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9909706546275394, f1=0.990990990990991,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004643334250431508
step: 10, loss: 0.000910637085326016
step: 20, loss: 0.0012163083301857114

```

```

step: 30, loss: 0.00020229211077094078
step: 40, loss: 0.0002528465411160141
step: 50, loss: 0.0004928504931740463
step: 60, loss: 0.0004011179553344846
step: 70, loss: 0.007199504878371954
step: 80, loss: 0.0034025602508336306
step: 90, loss: 0.0004216471570543945
step: 100, loss: 0.0554061233997345
step: 110, loss: 0.0023346422240138054
step: 120, loss: 0.0010591746540740132
step: 130, loss: 0.003094945102930069
step: 140, loss: 0.0012189242988824844
step: 150, loss: 0.004490633029490709
step: 160, loss: 0.0029848592821508646
step: 170, loss: 0.0005569394561462104
step: 180, loss: 0.04499557986855507
step: 190, loss: 0.001862545614130795
step: 200, loss: 0.0007465796661563218
step: 210, loss: 0.0007024129154160619
step: 220, loss: 0.00029567614546976984
step: 230, loss: 0.08257032930850983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9887892376681614, f1=0.9865771812080537,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004696170217357576
step: 10, loss: 0.0006795309600420296
step: 20, loss: 0.0008400175138376653
step: 30, loss: 0.021193798631429672
step: 40, loss: 0.00035588856553658843
step: 50, loss: 0.002888147020712495
step: 60, loss: 0.009091578423976898
step: 70, loss: 0.037951502948999405
step: 80, loss: 0.0002889778988901526
step: 90, loss: 7.363057375187054e-05
step: 100, loss: 0.00012154789146734402
step: 110, loss: 0.0005987868644297123
step: 120, loss: 9.280820086132735e-05
step: 130, loss: 9.357178350910544e-05
step: 140, loss: 0.0002481814008206129
step: 150, loss: 0.0017473198240622878
step: 160, loss: 0.0878690704703331
step: 170, loss: 0.0015980337047949433
step: 180, loss: 0.0035589993931353092

```

```

step: 190, loss: 0.0002089741756208241
step: 200, loss: 0.05193629115819931
step: 210, loss: 5.568963024416007e-05
step: 220, loss: 0.00014082467532716691
step: 230, loss: 0.00041995072388090193
/usr/local/lib/python3.7/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=0.9899441340782122, f1=0.9866071428571428,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019048460526391864
step: 10, loss: 0.0015877505065873265
step: 20, loss: 0.00015228136908262968
step: 30, loss: 0.002120697172358632
step: 40, loss: 0.00019299793348181993
step: 50, loss: 0.00011606172483880073
step: 60, loss: 0.00010267958714393899
step: 70, loss: 0.0001673384540481493
step: 80, loss: 0.0002528596669435501
step: 90, loss: 4.633162097888999e-05
step: 100, loss: 0.0010201395489275455
step: 110, loss: 0.05154578387737274
step: 120, loss: 0.0002976386749651283
step: 130, loss: 0.0009352618362754583
step: 140, loss: 0.0002564032911323011
step: 150, loss: 0.001875429879873991
step: 160, loss: 0.00040682547842152417
step: 170, loss: 0.0001553720358060673
step: 180, loss: 0.0018229311099275947
step: 190, loss: 0.004371581133455038
step: 200, loss: 9.367190068587661e-05
step: 210, loss: 0.000120550837891642
step: 220, loss: 0.00038920051883906126
step: 230, loss: 0.00042193764238618314
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932432432432432, f1=0.990990990990991,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.971002924023196e-05
step: 10, loss: 0.0002903903368860483
step: 20, loss: 0.00039863819256424904
step: 30, loss: 0.000286387512460351
step: 40, loss: 0.02188197895884514
step: 50, loss: 7.554612238891423e-05
step: 60, loss: 0.000121271688840352
step: 70, loss: 0.00013570513692684472
step: 80, loss: 0.00036409118911251426
step: 90, loss: 0.003626441117376089
step: 100, loss: 0.07196001708507538
step: 110, loss: 9.754667553352192e-05
step: 120, loss: 6.97956929798238e-05
step: 130, loss: 0.000757248664740473
step: 140, loss: 0.0001330306549789384
step: 150, loss: 0.001399158500134945
step: 160, loss: 0.0001438802428310737
step: 170, loss: 0.00021534078405238688
step: 180, loss: 0.002809458179399371
step: 190, loss: 0.00035020208451896906
step: 200, loss: 0.0016083671944215894
step: 210, loss: 0.0001393545389873907
step: 220, loss: 8.161953883245587e-05
step: 230, loss: 0.0001273989910259843
/usr/local/lib/python3.7/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=0.9899441340782122, f1=0.9866369710467707,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015924610488582402
step: 10, loss: 9.140477050095797e-05
step: 20, loss: 0.00029054583865217865
step: 30, loss: 0.00018266533152200282
step: 40, loss: 0.00011811235890490934
step: 50, loss: 5.8904413890559226e-05
step: 60, loss: 0.005658616777509451
step: 70, loss: 0.0001457331090932712
step: 80, loss: 0.00010465483501320705
step: 90, loss: 0.0012993935961276293
step: 100, loss: 8.507943857694045e-05
step: 110, loss: 0.00010212352935923263
step: 120, loss: 0.0011370356660336256
step: 130, loss: 0.0030976503621786833
step: 140, loss: 0.03446827828884125

```

```

step: 150, loss: 0.017429282888770103
step: 160, loss: 3.982575799454935e-05
step: 170, loss: 0.00010985361586790532
step: 180, loss: 0.00013630989997182041
step: 190, loss: 0.00677076680585742
step: 200, loss: 6.46244952804409e-05
step: 210, loss: 0.00012768343731295317
step: 220, loss: 6.674150063190609e-05
step: 230, loss: 0.0001840363984229043
/usr/local/lib/python3.7/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=0.9921787709497207, f1=0.9899216125419933,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.762962145032361e-05
step: 10, loss: 9.581191261531785e-05
step: 20, loss: 0.0004895448219031096
step: 30, loss: 0.02324669063091278
step: 40, loss: 0.0004614480712916702
step: 50, loss: 0.001743387896567583
step: 60, loss: 0.00010147656576009467
step: 70, loss: 0.0003213670279365033
step: 80, loss: 4.138280564802699e-05
step: 90, loss: 4.687644468504004e-05
step: 100, loss: 9.439212590223178e-05
step: 110, loss: 0.01814650371670723
step: 120, loss: 5.904030695091933e-05
step: 130, loss: 3.774951983359642e-05
step: 140, loss: 4.773293039761484e-05
step: 150, loss: 0.026236670091748238
step: 160, loss: 5.125538518768735e-05
step: 170, loss: 0.019134169444441795
step: 180, loss: 5.317868271959014e-05
step: 190, loss: 5.281284393277019e-05
step: 200, loss: 0.0001443083892809227
step: 210, loss: 4.368794179754332e-05
step: 220, loss: 5.249290552455932e-05
step: 230, loss: 0.0002761329524219036
/usr/local/lib/python3.7/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=0.9899216125419933, f1=0.987736900780379,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.585271355812438e-05
step: 10, loss: 4.559478838928044e-05
step: 20, loss: 5.4429980082204565e-05
step: 30, loss: 5.617055285256356e-05
step: 40, loss: 9.446749754715711e-05
step: 50, loss: 0.00010706916509661824
step: 60, loss: 0.0017604012973606586
step: 70, loss: 0.0001124758564401418
step: 80, loss: 6.843705341452733e-05
step: 90, loss: 4.01680517825298e-05
step: 100, loss: 3.5370478144614026e-05
step: 110, loss: 5.4771688155597076e-05
step: 120, loss: 6.483416655100882e-05
step: 130, loss: 4.034355515614152e-05
step: 140, loss: 0.010269710794091225
step: 150, loss: 0.0015583318891003728
step: 160, loss: 4.77635367133189e-05
step: 170, loss: 9.354069334222004e-05
step: 180, loss: 0.0574229471385479
step: 190, loss: 0.00012736656935885549
step: 200, loss: 2.8747303076670505e-05
step: 210, loss: 0.03985742852091789
step: 220, loss: 5.35196959390305e-05
step: 230, loss: 0.032320909202098846
/usr/local/lib/python3.7/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=0.9899216125419933, f1=0.9866071428571428,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015366336447186768
step: 10, loss: 0.00028588209534063935
step: 20, loss: 0.00017949902394320816
step: 30, loss: 0.0005673299892805517
step: 40, loss: 6.58846547594294e-05
step: 50, loss: 0.0010856824228540063
step: 60, loss: 4.973892282578163e-05
step: 70, loss: 3.365257725818083e-05
step: 80, loss: 3.9988855860428885e-05
step: 90, loss: 8.420318044954911e-05
step: 100, loss: 2.6825258828466758e-05
step: 110, loss: 0.023835834115743637
step: 120, loss: 0.0287881251424551
step: 130, loss: 5.7667286455398425e-05
step: 140, loss: 4.485772660700604e-05

```



```

step: 150, loss: 6.394975207513198e-05
step: 160, loss: 0.0003004639584105462
step: 170, loss: 4.8339195927837864e-05
step: 180, loss: 5.0767845095833763e-05
step: 190, loss: 0.00016646440781187266
step: 200, loss: 0.00012526678619906306
step: 210, loss: 2.9432179871946573e-05
step: 220, loss: 4.2044484871439636e-05
step: 230, loss: 5.625276389764622e-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=0.9899441340782122, f1=0.9866369710467707,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.658117930172011e-05
step: 10, loss: 0.022314593195915222
step: 20, loss: 0.0003347222809679806
step: 30, loss: 5.451887409435585e-05
step: 40, loss: 0.016758298501372337
step: 50, loss: 4.753984831040725e-05
step: 60, loss: 3.249845758546144e-05
step: 70, loss: 4.219771653879434e-05
step: 80, loss: 4.65030170744285e-05
step: 90, loss: 3.777271558647044e-05
step: 100, loss: 3.192111398675479e-05
step: 110, loss: 0.00011006835848093033
step: 120, loss: 2.2217247533262707e-05
step: 130, loss: 3.960561298299581e-05
step: 140, loss: 0.0001905227400129661
step: 150, loss: 3.2934123737504706e-05
step: 160, loss: 0.0013870032271370292
step: 170, loss: 3.2782678317744285e-05
step: 180, loss: 4.6477267460431904e-05
step: 190, loss: 0.00010860917973332107
step: 200, loss: 3.3227737731067464e-05
step: 210, loss: 5.059716568212025e-05
step: 220, loss: 8.927836461225525e-05
step: 230, loss: 4.1355353459948674e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9899216125419933, f1=0.9866071428571428,
best_f1=0.990990990990991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.556605431891512e-05
step: 10, loss: 2.1397576347226277e-05
step: 20, loss: 4.1050770960282534e-05
step: 30, loss: 4.5117936679162085e-05
step: 40, loss: 0.00010162361286347732
step: 50, loss: 0.037615664303302765
step: 60, loss: 7.692714279983193e-05
step: 70, loss: 4.19476637034677e-05
step: 80, loss: 7.360973540926352e-05
step: 90, loss: 5.662188777932897e-05
step: 100, loss: 3.5020028008148074e-05
step: 110, loss: 3.059484879486263e-05
step: 120, loss: 6.0382280935300514e-05
step: 130, loss: 0.013226364739239216
step: 140, loss: 2.919064718298614e-05
step: 150, loss: 8.845584670780227e-05
step: 160, loss: 3.628895865404047e-05
step: 170, loss: 2.6504643756197765e-05
step: 180, loss: 0.00014236057177186012
step: 190, loss: 4.08197010983713e-05
step: 200, loss: 0.00011281197657808661
step: 210, loss: 3.8749036320950836e-05
step: 220, loss: 0.00010953253513434902
step: 230, loss: 3.220771031919867e-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=0.9899216125419933, f1=0.9866071428571428,
best_f1=0.990990990990991
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:14, 171.33it/s]
load_f1 = 0.9921612541993281
real_f1 = 0.9910514541387023
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, 194.15it/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: 0.5953010320663452
```

```
step: 10, loss: 0.5200695991516113
```

```
step: 20, loss: 0.3919454514980316
```

step: 30, loss: 0.06429942697286606
step: 40, loss: 0.12061807513237
step: 50, loss: 0.07541155815124512
step: 60, loss: 0.03590308129787445
step: 70, loss: 0.06819230318069458
step: 80, loss: 0.06723489612340927
step: 90, loss: 0.1045457124710083
step: 100, loss: 0.08376342058181763
step: 110, loss: 0.05678074434399605
step: 120, loss: 0.1623467355966568
step: 130, loss: 0.08345045894384384
step: 140, loss: 0.10968698561191559
step: 150, loss: 0.013597307726740837
step: 160, loss: 0.013697321526706219
step: 170, loss: 0.1577625572681427
step: 180, loss: 0.015797702595591545
step: 190, loss: 0.010125957429409027
step: 200, loss: 0.20648911595344543
step: 210, loss: 0.11565383523702621
step: 220, loss: 0.16624142229557037
step: 230, loss: 0.14090774953365326
step: 240, loss: 0.04998905584216118
step: 250, loss: 0.011450892314314842
step: 260, loss: 0.049163706600666046
step: 270, loss: 0.0077204336412250996
step: 280, loss: 0.015693912282586098
step: 290, loss: 0.0748322606086731
step: 300, loss: 0.053585540503263474
step: 310, loss: 0.16877290606498718
step: 320, loss: 0.09099777787923813
step: 330, loss: 0.01520766131579876
step: 340, loss: 0.05707579106092453
step: 350, loss: 0.0341101735830307
step: 360, loss: 0.014098435640335083
step: 370, loss: 0.08664780110120773
step: 380, loss: 0.0056077903136610985
step: 390, loss: 0.17099414765834808
step: 400, loss: 0.1805538684129715
step: 410, loss: 0.048218633979558945
step: 420, loss: 0.017758561298251152
step: 430, loss: 0.16367316246032715
step: 440, loss: 0.010059033520519733
step: 450, loss: 0.002071286551654339
step: 460, loss: 0.015454111620783806
step: 470, loss: 0.21628232300281525
step: 480, loss: 0.06184372678399086
step: 490, loss: 0.07559242099523544
step: 500, loss: 0.06607841700315475

```

step: 510, loss: 0.014942812733352184
step: 520, loss: 0.08155977725982666
step: 530, loss: 0.0027095074765384197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9483870967741935, f1=0.9382488479262673,
best_f1=0.9382488479262673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.154183492064476
step: 10, loss: 0.049660634249448776
step: 20, loss: 0.02880674973130226
step: 30, loss: 0.059334106743335724
step: 40, loss: 0.1268710345029831
step: 50, loss: 0.11516132950782776
step: 60, loss: 0.04892309010028839
step: 70, loss: 0.03509913757443428
step: 80, loss: 0.02029813826084137
step: 90, loss: 0.00857858918607235
step: 100, loss: 0.004670572932809591
step: 110, loss: 0.016221223399043083
step: 120, loss: 0.03373376652598381
step: 130, loss: 0.037467729300260544
step: 140, loss: 0.02714325301349163
step: 150, loss: 0.01005780603736639
step: 160, loss: 0.003365442855283618
step: 170, loss: 0.040678899735212326
step: 180, loss: 0.12810944020748138
step: 190, loss: 0.06405968964099884
step: 200, loss: 0.004800031892955303
step: 210, loss: 0.036824632436037064
step: 220, loss: 0.04744550585746765
step: 230, loss: 0.006193278357386589
step: 240, loss: 0.04181142523884773
step: 250, loss: 0.0332406722009182
step: 260, loss: 0.010265486314892769
step: 270, loss: 0.1316649168729782
step: 280, loss: 0.031405653804540634
step: 290, loss: 0.007974883541464806
step: 300, loss: 0.16684935986995697
step: 310, loss: 0.007488710805773735
step: 320, loss: 0.0916672870516777

```

```

step: 330, loss: 0.04665636643767357
step: 340, loss: 0.011675097979605198
step: 350, loss: 0.0009047830826602876
step: 360, loss: 0.029645860195159912
step: 370, loss: 0.06380616873502731
step: 380, loss: 0.03876405209302902
step: 390, loss: 0.042053673416376114
step: 400, loss: 0.044280052185058594
step: 410, loss: 0.032233476638793945
step: 420, loss: 0.13628354668617249
step: 430, loss: 0.005426608491688967
step: 440, loss: 0.20175153017044067
step: 450, loss: 0.010247398167848587
step: 460, loss: 0.024041656404733658
step: 470, loss: 0.04255792126059532
step: 480, loss: 0.3116149306297302
step: 490, loss: 0.016037048771977425
step: 500, loss: 0.1704459935426712
step: 510, loss: 0.006460798904299736
step: 520, loss: 0.06653989106416702
step: 530, loss: 0.003907320089638233
/usr/local/lib/python3.7/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=0.9472701819878675, f1=0.9431345353675451,
best_f1=0.9382488479262673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05027144402265549
step: 10, loss: 0.04008779674768448
step: 20, loss: 0.14028871059417725
step: 30, loss: 0.07502831518650055
step: 40, loss: 0.06435532122850418
step: 50, loss: 0.01651102490723133
step: 60, loss: 0.00724867545068264
step: 70, loss: 0.0031483054626733065
step: 80, loss: 0.001659761997871101
step: 90, loss: 0.006048875395208597
step: 100, loss: 0.04176735132932663
step: 110, loss: 0.002508860547095537
step: 120, loss: 0.0030644419603049755
step: 130, loss: 0.0012291207676753402
step: 140, loss: 0.08423801511526108
step: 150, loss: 0.016563549637794495
step: 160, loss: 0.020215043798089027
step: 170, loss: 0.042252540588378906
step: 180, loss: 0.006097907200455666

```

```

step: 190, loss: 0.0055914330296218395
step: 200, loss: 0.04359522834420204
step: 210, loss: 0.11947059631347656
step: 220, loss: 0.016167327761650085
step: 230, loss: 0.13185176253318787
step: 240, loss: 0.004450679291039705
step: 250, loss: 0.027493776753544807
step: 260, loss: 0.009249014779925346
step: 270, loss: 0.0008830113802105188
step: 280, loss: 0.231331467628479
step: 290, loss: 0.006610377691686153
step: 300, loss: 0.026521841064095497
step: 310, loss: 0.046905990689992905
step: 320, loss: 0.018393544480204582
step: 330, loss: 0.0015515376580879092
step: 340, loss: 0.004632241558283567
step: 350, loss: 0.0017612148076295853
step: 360, loss: 0.03758009523153305
step: 370, loss: 0.009387371130287647
step: 380, loss: 0.003980901557952166
step: 390, loss: 0.02337092161178589
step: 400, loss: 0.01533243153244257
step: 410, loss: 0.01840922050178051
step: 420, loss: 0.11250169575214386
step: 430, loss: 0.01341501995921135
step: 440, loss: 0.013290584087371826
step: 450, loss: 0.1057419627904892
step: 460, loss: 0.04370654746890068
step: 470, loss: 0.06232385337352753
step: 480, loss: 0.0038946785498410463
step: 490, loss: 0.005571685265749693
step: 500, loss: 0.009645103476941586
step: 510, loss: 0.002117888303473592
step: 520, loss: 0.06212342530488968
step: 530, loss: 0.12532402575016022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.958217270194986, f1=0.9458583988894032,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00811837986111641

```


step: 10, loss: 0.0013886063825339079
step: 20, loss: 0.0012892571976408362
step: 30, loss: 0.008222454227507114
step: 40, loss: 0.002195967361330986
step: 50, loss: 0.016440000385046005
step: 60, loss: 0.00834150705486536
step: 70, loss: 0.0072612883523106575
step: 80, loss: 0.0012215941678732634
step: 90, loss: 0.09089246392250061
step: 100, loss: 0.00482352077960968
step: 110, loss: 0.00977187417447567
step: 120, loss: 0.0013633803464472294
step: 130, loss: 0.0038272978272289038
step: 140, loss: 0.001843369216658175
step: 150, loss: 0.011442362330853939
step: 160, loss: 0.016336631029844284
step: 170, loss: 0.0025898481253534555
step: 180, loss: 0.00036036476376466453
step: 190, loss: 0.002142041688784957
step: 200, loss: 0.0009048093925230205
step: 210, loss: 0.01576019637286663
step: 220, loss: 0.004767796490341425
step: 230, loss: 0.0737207680940628
step: 240, loss: 0.012015841901302338
step: 250, loss: 0.00952798593789339
step: 260, loss: 0.0031532992143183947
step: 270, loss: 0.001727303140796721
step: 280, loss: 0.034505993127822876
step: 290, loss: 0.027221867814660072
step: 300, loss: 0.00031099896295927465
step: 310, loss: 0.005779766943305731
step: 320, loss: 0.026607759296894073
step: 330, loss: 0.002452367451041937
step: 340, loss: 0.0011085720034316182
step: 350, loss: 0.004364626482129097
step: 360, loss: 0.003305225633084774
step: 370, loss: 0.003833555383607745
step: 380, loss: 0.01696093939244747
step: 390, loss: 0.000828360382001847
step: 400, loss: 0.0010118790669366717
step: 410, loss: 0.0009796704398468137
step: 420, loss: 0.06471575051546097
step: 430, loss: 0.004129515960812569
step: 440, loss: 0.005794293247163296
step: 450, loss: 0.0016518462216481566
step: 460, loss: 0.001261685392819345
step: 470, loss: 0.002913418924435973
step: 480, loss: 0.11133081465959549

```
step: 490, loss: 0.004838601220399141
step: 500, loss: 0.0034816621337085962
step: 510, loss: 0.016770079731941223
step: 520, loss: 0.07445119321346283
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 530, loss: 0.008643344976007938
/usr/local/lib/python3.7/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=0.9481961147086033, f1=0.9419414770088249,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08672703802585602
step: 10, loss: 0.03994191810488701
step: 20, loss: 0.001502394792623818
step: 30, loss: 0.00034056566073559225
step: 40, loss: 0.0071219573728740215
step: 50, loss: 0.00044054246973246336
step: 60, loss: 0.058393195271492004
step: 70, loss: 0.00510170916095376
step: 80, loss: 0.001150303054600954
step: 90, loss: 0.001458090147934854
step: 100, loss: 0.018243182450532913
step: 110, loss: 0.0003037918359041214
step: 120, loss: 0.018332138657569885
step: 130, loss: 0.00014067665324546397
step: 140, loss: 0.0003951630787923932
step: 150, loss: 0.010430952534079552
step: 160, loss: 0.0001793931151041761
step: 170, loss: 0.01739520952105522
step: 180, loss: 0.0006840730202384293
step: 190, loss: 0.00011686503421515226
step: 200, loss: 0.00023819509078748524
step: 210, loss: 0.02816096507012844
step: 220, loss: 0.00026777261518873274
step: 230, loss: 0.0030608256347477436
step: 240, loss: 0.01978844404220581
step: 250, loss: 0.06621238589286804
step: 260, loss: 0.016939600929617882
step: 270, loss: 0.004877133294939995
step: 280, loss: 0.014897503890097141
step: 290, loss: 0.1316748559474945
step: 300, loss: 0.01220589317381382
step: 310, loss: 0.05167291685938835
step: 320, loss: 0.10881582647562027
step: 330, loss: 0.0029724598862230778
```

```
step: 340, loss: 0.0033841687254607677
step: 350, loss: 0.004631441552191973
step: 360, loss: 0.004305220674723387
step: 370, loss: 0.012677911669015884
step: 380, loss: 0.0002825965639203787
step: 390, loss: 0.0005456332000903785
step: 400, loss: 0.004678497556596994
step: 410, loss: 0.0029591647908091545
step: 420, loss: 0.0020551502238959074
step: 430, loss: 0.01967567391693592
step: 440, loss: 0.08412374556064606
step: 450, loss: 0.0015359454555436969
step: 460, loss: 0.001439766027033329
step: 470, loss: 0.011971416883170605
step: 480, loss: 0.0064429957419633865
step: 490, loss: 0.007675500586628914
step: 500, loss: 0.007157857995480299
step: 510, loss: 0.0214249175041914
step: 520, loss: 0.00690931361168623
step: 530, loss: 0.005327304825186729
/usr/local/lib/python3.7/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=0.9511970534069981, f1=0.9476102941176471,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011296680895611644
step: 10, loss: 0.00045909034088253975
step: 20, loss: 0.1021212786436081
step: 30, loss: 0.0014521169941872358
step: 40, loss: 0.003854612819850445
step: 50, loss: 0.017428262159228325
step: 60, loss: 0.00041303157922811806
step: 70, loss: 0.0003201634099241346
step: 80, loss: 0.010188540443778038
step: 90, loss: 0.000153069689986296
step: 100, loss: 0.0007429805118590593
step: 110, loss: 0.0064664254896342754
step: 120, loss: 0.00680412957444787
step: 130, loss: 0.000556844926904887
step: 140, loss: 0.008547293953597546
step: 150, loss: 0.00021002639550715685
step: 160, loss: 0.0005513591459020972
step: 170, loss: 0.00017833524907473475
step: 180, loss: 0.0024317638017237186
step: 190, loss: 0.0003459958534222096
```

```

step: 200, loss: 0.00019543161033652723
step: 210, loss: 0.0008182215387932956
step: 220, loss: 0.0002021320105995983
step: 230, loss: 0.001567895757034421
step: 240, loss: 0.0009608220425434411
step: 250, loss: 0.0023560530971735716
step: 260, loss: 0.00012662261724472046
step: 270, loss: 0.057864960283041
step: 280, loss: 0.0009782634442672133
step: 290, loss: 0.0008924663998186588
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 300, loss: 0.004382743034511805
step: 310, loss: 0.0004576299397740513
step: 320, loss: 0.0004563039110507816
step: 330, loss: 0.000365561485523358
step: 340, loss: 0.03948380798101425
step: 350, loss: 0.008149301633238792
step: 360, loss: 0.035606835037469864
step: 370, loss: 0.007954669184982777
step: 380, loss: 0.0004322724125813693
step: 390, loss: 0.2099991738796234
step: 400, loss: 0.0003808906767517328
step: 410, loss: 0.0026982619892805815
step: 420, loss: 0.0006531385006383061
step: 430, loss: 0.00022124650422483683
step: 440, loss: 0.00370218837633729
step: 450, loss: 0.03791595995426178
step: 460, loss: 0.00019199909002054483
step: 470, loss: 0.013584228232502937
step: 480, loss: 0.005730025935918093
step: 490, loss: 0.011721561662852764
step: 500, loss: 0.008019937202334404
step: 510, loss: 0.0036728139966726303
step: 520, loss: 0.011282071471214294
step: 530, loss: 0.06972137093544006
/usr/local/lib/python3.7/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=0.9451360073766714, f1=0.9460083064143978,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009415716049261391
step: 10, loss: 0.002325089182704687
step: 20, loss: 0.0008668775553815067
step: 30, loss: 0.005759293679147959
step: 40, loss: 0.002465000143274665

```

step: 50, loss: 0.02351567894220352
step: 60, loss: 0.0014329884434118867
step: 70, loss: 0.004387340042740107
step: 80, loss: 0.005621941760182381
step: 90, loss: 0.0002529052144382149
step: 100, loss: 0.0003777820966206491
step: 110, loss: 0.000426614802563563
step: 120, loss: 0.0014657345600426197
step: 130, loss: 0.0010181607212871313
step: 140, loss: 0.0009584350627847016
step: 150, loss: 0.0007973283645696938
step: 160, loss: 0.0003692817408591509
step: 170, loss: 9.454281098442152e-05
step: 180, loss: 0.0004700772406067699
step: 190, loss: 0.0028127222321927547
step: 200, loss: 0.00079909194028005
step: 210, loss: 0.0022963504306972027
step: 220, loss: 0.0005118183325976133
step: 230, loss: 0.01987779513001442
step: 240, loss: 0.00010059119813377038
step: 250, loss: 0.003056284273043275
step: 260, loss: 0.00010383346671005711
step: 270, loss: 0.0009555871365591884
step: 280, loss: 0.0001535707269795239
step: 290, loss: 6.871300865896046e-05
step: 300, loss: 0.004774641711264849
step: 310, loss: 8.677273581270128e-05
step: 320, loss: 4.422032361617312e-05
step: 330, loss: 0.012065333314239979
step: 340, loss: 0.015702642500400543
step: 350, loss: 6.558381573995575e-05
step: 360, loss: 0.0009854478994384408
step: 370, loss: 5.4143361921887845e-05
step: 380, loss: 7.875442679505795e-05
step: 390, loss: 0.25651639699935913
step: 400, loss: 0.023224903270602226
step: 410, loss: 0.0035689296200871468
step: 420, loss: 0.00837146956473589
step: 430, loss: 9.975626016966999e-05
step: 440, loss: 0.10177874565124512
step: 450, loss: 0.003297198098152876
step: 460, loss: 0.0013058248441666365
step: 470, loss: 0.017204638570547104
step: 480, loss: 0.1966889649629593
step: 490, loss: 0.0005845110863447189
step: 500, loss: 0.0042470525950193405
step: 510, loss: 0.0018186442321166396
step: 520, loss: 0.027282409369945526

```

step: 530, loss: 0.013668522238731384
/usr/local/lib/python3.7/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=0.952073732718894, f1=0.9472718936267768,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003421911096666008
step: 10, loss: 0.04430776834487915
step: 20, loss: 0.00018301029922440648
step: 30, loss: 0.000357154494849965
step: 40, loss: 0.00011405138502595946
step: 50, loss: 0.0010857112938538194
step: 60, loss: 0.004523130599409342
step: 70, loss: 0.0002477738307788968
step: 80, loss: 0.05202683433890343
step: 90, loss: 0.009683400392532349
step: 100, loss: 0.008157510310411453
step: 110, loss: 0.0001115069244406186
step: 120, loss: 0.015726476907730103
step: 130, loss: 0.00029648246709257364
step: 140, loss: 0.008960671722888947
step: 150, loss: 0.0002359763311687857
step: 160, loss: 0.0005701487534679472
step: 170, loss: 0.001454771962016821
step: 180, loss: 0.00025981859653256834
step: 190, loss: 0.0021852299105376005
step: 200, loss: 0.001932229264639318
step: 210, loss: 0.0010945205576717854
step: 220, loss: 0.004095542710274458
step: 230, loss: 0.0046088723465800285
step: 240, loss: 7.805214409017935e-05
step: 250, loss: 6.059719453332946e-05
step: 260, loss: 0.0010117841884493828
step: 270, loss: 0.001239991164766252
step: 280, loss: 0.11883196979761124
step: 290, loss: 0.00029556575464084744
step: 300, loss: 0.0002518379769753665
step: 310, loss: 0.00029686346533708274
step: 320, loss: 0.0034416439011693
step: 330, loss: 0.00037116167368367314
step: 340, loss: 0.001141330343671143
step: 350, loss: 0.00011068614549003541
step: 360, loss: 0.0027144495397806168
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 370, loss: 0.03875020891427994

```

```

step: 380, loss: 0.0005909426836296916
step: 390, loss: 0.13747072219848633
step: 400, loss: 9.694777691038325e-05
step: 410, loss: 0.00019833717669826
step: 420, loss: 0.003280828008428216
step: 430, loss: 0.027667630463838577
step: 440, loss: 0.001952814287506044
step: 450, loss: 0.00011622800957411528
step: 460, loss: 0.008393729105591774
step: 470, loss: 0.00017025170382112265
step: 480, loss: 6.465730984928086e-05
step: 490, loss: 0.0011014469200745225
step: 500, loss: 0.00032130329054780304
step: 510, loss: 7.834864663891494e-05
step: 520, loss: 0.002064620843157172
step: 530, loss: 0.0001784974883776158
/usr/local/lib/python3.7/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=0.9552376557452701, f1=0.9451360073766714,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00151829665992409
step: 10, loss: 5.984495874145068e-05
step: 20, loss: 0.0017430844018235803
step: 30, loss: 0.000382820435333997
step: 40, loss: 0.0005024370620958507
step: 50, loss: 4.548121432890184e-05
step: 60, loss: 0.00011830421863123775
step: 70, loss: 0.0011385661782696843
step: 80, loss: 0.009637204930186272
step: 90, loss: 7.667236786801368e-05
step: 100, loss: 0.00011059609096264467
step: 110, loss: 3.516943979775533e-05
step: 120, loss: 3.2334606657968834e-05
step: 130, loss: 0.01787310466170311
step: 140, loss: 0.00022828514920547605
step: 150, loss: 0.0011172662489116192
step: 160, loss: 0.001829368993639946
step: 170, loss: 0.004351183772087097
step: 180, loss: 0.00019235158106312156
step: 190, loss: 6.280694651650265e-05
step: 200, loss: 0.00016631008475087583
step: 210, loss: 0.0001316045963903889
step: 220, loss: 9.460585715714842e-05
step: 230, loss: 5.1706043450394645e-05

```

```

step: 240, loss: 9.056620910996571e-05
step: 250, loss: 0.00038061095983721316
step: 260, loss: 0.027122292667627335
step: 270, loss: 2.3930768293212168e-05
step: 280, loss: 0.0991809293627739
step: 290, loss: 0.0359790101647377
step: 300, loss: 0.0009327777079306543
step: 310, loss: 0.027077531442046165
step: 320, loss: 0.0005519681726582348
step: 330, loss: 0.00021837037638761103
step: 340, loss: 0.00010847513476619497
step: 350, loss: 0.00817265547811985
step: 360, loss: 6.33819290669635e-05
step: 370, loss: 0.009085147641599178
step: 380, loss: 0.0003762776032090187
step: 390, loss: 0.00036509957863017917
step: 400, loss: 0.016375549137592316
step: 410, loss: 4.799775342689827e-05
step: 420, loss: 3.27370798913762e-05
step: 430, loss: 4.863398498855531e-05
step: 440, loss: 0.001666210126131773
step: 450, loss: 6.391402712324634e-05
step: 460, loss: 0.00019159403746016324
step: 470, loss: 2.8251688490854576e-05
step: 480, loss: 0.0001819846365833655
step: 490, loss: 0.0001939307403517887
step: 500, loss: 0.002631369512528181
step: 510, loss: 0.00015676797193009406
step: 520, loss: 4.959703073836863e-05
step: 530, loss: 0.000287866365397349
/usr/local/lib/python3.7/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=0.9483960948396094, f1=0.9423791821561338,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004516872111707926
step: 10, loss: 3.691033271024935e-05
step: 20, loss: 2.568901254562661e-05
step: 30, loss: 4.078530764672905e-05
step: 40, loss: 8.327847172040492e-05
step: 50, loss: 0.0025866262149065733
step: 60, loss: 4.789245576830581e-05
step: 70, loss: 0.00011298748722765595
step: 80, loss: 5.304690057528205e-05
step: 90, loss: 4.148250081925653e-05

```



```
step: 100, loss: 3.395804014871828e-05
step: 110, loss: 1.9341418010299094e-05
step: 120, loss: 4.6189506974769756e-05
step: 130, loss: 4.6451670641545206e-05
step: 140, loss: 0.009343519806861877
step: 150, loss: 8.572271326556802e-05
step: 160, loss: 8.14757077023387e-05
step: 170, loss: 8.246200741268694e-05
step: 180, loss: 6.42090235487558e-05
step: 190, loss: 5.902381963096559e-05
step: 200, loss: 0.00020854327885899693
step: 210, loss: 0.0005720116896554828
step: 220, loss: 0.0012190780835226178
step: 230, loss: 0.0006142252823337913
step: 240, loss: 0.00012081763270543888
step: 250, loss: 3.203244705218822e-05
step: 260, loss: 0.0007070240098983049
step: 270, loss: 0.0006315024802461267
step: 280, loss: 0.00022930702834855765
step: 290, loss: 3.727732109837234e-05
step: 300, loss: 3.171213757013902e-05
step: 310, loss: 3.089652818744071e-05
step: 320, loss: 0.001349269994534552
step: 330, loss: 0.0007900111377239227
step: 340, loss: 0.00016103865345939994
step: 350, loss: 3.255403498769738e-05
step: 360, loss: 0.0001027253019856289
step: 370, loss: 0.0038470595609396696
step: 380, loss: 0.0012119374005123973
step: 390, loss: 0.0006413719383999705
step: 400, loss: 0.002413311507552862
step: 410, loss: 0.0020762921776622534
step: 420, loss: 0.0001921385119203478
step: 430, loss: 2.2444453861680813e-05
step: 440, loss: 0.00046839501010254025
step: 450, loss: 8.11407808214426e-05
step: 460, loss: 3.929471131414175e-05
step: 470, loss: 6.117042357800528e-05
step: 480, loss: 5.05584612255916e-05
step: 490, loss: 0.0002433390764053911
step: 500, loss: 0.1056264340877533
step: 510, loss: 0.00020959123503416777
step: 520, loss: 0.000959648285061121
step: 530, loss: 0.0021911554504185915
/usr/local/lib/python3.7/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=0.9515905947441217, f1=0.9463056447911886,
```

```
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006322403904050589
step: 10, loss: 0.0015602257335558534
step: 20, loss: 0.00011872783943545073
step: 30, loss: 0.015393568202853203
step: 40, loss: 0.00012109379167668521
step: 50, loss: 0.0005115606472827494
step: 60, loss: 0.03509936481714249
step: 70, loss: 0.00010749335342552513
step: 80, loss: 0.0015397254610434175
step: 90, loss: 0.0021252124570310116
step: 100, loss: 0.0003230192814953625
step: 110, loss: 0.0002496186352800578
step: 120, loss: 0.0006458996795117855
step: 130, loss: 6.478155410150066e-05
step: 140, loss: 0.0002510866615921259
step: 150, loss: 3.65949199476745e-05
step: 160, loss: 4.132201866013929e-05
step: 170, loss: 8.601608715252951e-05
step: 180, loss: 0.00016470788978040218
step: 190, loss: 0.001307033933699131
step: 200, loss: 0.004047782626003027
step: 210, loss: 0.00020714988932013512
step: 220, loss: 0.0027787787839770317
step: 230, loss: 2.971925823658239e-05
step: 240, loss: 0.0009442750015296042
step: 250, loss: 0.0002472131745889783
step: 260, loss: 3.75658055418171e-05
step: 270, loss: 0.005881187971681356
step: 280, loss: 4.9141282943310216e-05
step: 290, loss: 0.018731269985437393
step: 300, loss: 5.797547419206239e-05
step: 310, loss: 5.9982903621857986e-05
step: 320, loss: 4.878996332990937e-05
step: 330, loss: 0.00010187848238274455
step: 340, loss: 5.5737520597176626e-05
step: 350, loss: 0.00018547693616710603
step: 360, loss: 3.8943366234889254e-05
step: 370, loss: 0.00037264436832629144
step: 380, loss: 4.027924660476856e-05
step: 390, loss: 6.823305739089847e-05
step: 400, loss: 4.2416115320520476e-05
step: 410, loss: 5.1845068810507655e-05
step: 420, loss: 0.00010004961950471625
step: 430, loss: 0.0019107777625322342
```

```

step: 440, loss: 0.0002374912437517196
step: 450, loss: 0.00047856458695605397
step: 460, loss: 0.00025412815739400685
step: 470, loss: 0.00015654573508072644
step: 480, loss: 0.001248369924724102
step: 490, loss: 0.0001046025863615796
step: 500, loss: 8.484769205097109e-05
step: 510, loss: 5.9543886891333386e-05
step: 520, loss: 4.70700251753442e-05
step: 530, loss: 2.8769551136065274e-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=0.9516728624535317, f1=0.9483960948396094,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.508215351961553e-05
step: 10, loss: 3.25494802382309e-05
step: 20, loss: 8.422709652222693e-05
step: 30, loss: 0.001996537670493126
step: 40, loss: 0.00036912443465553224
step: 50, loss: 0.05213826149702072
step: 60, loss: 0.002316227415576577
step: 70, loss: 5.384874020819552e-05
step: 80, loss: 3.4743115975288674e-05
step: 90, loss: 6.851986836409196e-05
step: 100, loss: 0.00011632305540842935
step: 110, loss: 4.570989040075801e-05
step: 120, loss: 2.035091347352136e-05
step: 130, loss: 0.0006812595529481769
step: 140, loss: 3.043460674234666e-05
step: 150, loss: 2.0853796740993857e-05
step: 160, loss: 2.2842878024675883e-05
step: 170, loss: 5.225803397479467e-05
step: 180, loss: 0.04732867702841759
step: 190, loss: 0.00012119835446355864
step: 200, loss: 5.707847230951302e-05
step: 210, loss: 0.0032270594965666533
step: 220, loss: 0.0009411857463419437
step: 230, loss: 9.856878023128957e-05
step: 240, loss: 0.0004342690226621926
step: 250, loss: 5.314718873705715e-05
step: 260, loss: 5.042640623287298e-05
step: 270, loss: 0.004396738018840551
step: 280, loss: 0.0012514840345829725
step: 290, loss: 0.0012762723490595818

```

```

step: 300, loss: 0.0010474640876054764
step: 310, loss: 6.35500400676392e-05
step: 320, loss: 2.4749740987317637e-05
step: 330, loss: 7.93009385233745e-05
step: 340, loss: 0.00010095077595906332
step: 350, loss: 0.00013093584857415408
step: 360, loss: 0.02130068466067314
step: 370, loss: 4.4449003326008096e-05
step: 380, loss: 0.00020458968356251717
step: 390, loss: 9.691878949524835e-05
step: 400, loss: 4.070406794198789e-05
step: 410, loss: 0.0035414414014667273
step: 420, loss: 0.008522282354533672
step: 430, loss: 0.0165618397295475
step: 440, loss: 0.00023606308968737721
step: 450, loss: 0.00010063411173177883
step: 460, loss: 2.722367935348302e-05
step: 470, loss: 4.6299450332298875e-05
step: 480, loss: 0.0002060380211332813
step: 490, loss: 0.00046537304297089577
step: 500, loss: 8.029999298742041e-05
step: 510, loss: 0.00038576655788347125
step: 520, loss: 2.361046426813118e-05
step: 530, loss: 3.583168290788308e-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=0.9530013959981387, f1=0.9471243042671614,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002899735001847148
step: 10, loss: 3.725324495462701e-05
step: 20, loss: 1.5705638361396268e-05
step: 30, loss: 5.3411797125590965e-05
step: 40, loss: 2.022801163548138e-05
step: 50, loss: 0.0002803613606374711
step: 60, loss: 3.405364623176865e-05
step: 70, loss: 3.156448292429559e-05
step: 80, loss: 2.1430920241982676e-05
step: 90, loss: 0.03544510155916214
step: 100, loss: 0.0010257114190608263
step: 110, loss: 0.0038647844921797514
step: 120, loss: 4.44190991402138e-05
step: 130, loss: 7.068474224070087e-05
step: 140, loss: 3.496090721455403e-05
step: 150, loss: 8.630276715848595e-05

```

```

step: 160, loss: 4.622164851753041e-05
step: 170, loss: 3.592195571400225e-05
step: 180, loss: 9.663825767347589e-05
step: 190, loss: 4.940460712532513e-05
step: 200, loss: 5.4963300499366596e-05
step: 210, loss: 3.1042676710058004e-05
step: 220, loss: 0.00016325051547028124
step: 230, loss: 2.3452907043974847e-05
step: 240, loss: 5.376594708650373e-05
step: 250, loss: 1.858145697042346e-05
step: 260, loss: 1.991870340134483e-05
step: 270, loss: 1.828344102250412e-05
step: 280, loss: 3.766911686398089e-05
step: 290, loss: 5.594380127149634e-05
step: 300, loss: 2.3889768272056244e-05
step: 310, loss: 0.0008751049754209816
step: 320, loss: 0.00165723473764956
step: 330, loss: 0.0031722020357847214
step: 340, loss: 4.132400135858916e-05
step: 350, loss: 1.946805787156336e-05
step: 360, loss: 1.7277674487559125e-05
step: 370, loss: 3.3465719752712175e-05
step: 380, loss: 3.6147390346741304e-05
step: 390, loss: 0.008200565353035927
step: 400, loss: 2.637370016600471e-05
step: 410, loss: 0.002237220061942935
step: 420, loss: 1.740433435770683e-05
step: 430, loss: 0.007871371693909168
step: 440, loss: 2.8821066734963097e-05
step: 450, loss: 2.1826102965860628e-05
step: 460, loss: 0.001180828083306551
step: 470, loss: 1.6689038602635264e-05
step: 480, loss: 4.484578312258236e-05
step: 490, loss: 6.53576644253917e-05
step: 500, loss: 4.3657710193656385e-05
step: 510, loss: 2.234743078588508e-05
step: 520, loss: 0.0004579833475872874
step: 530, loss: 2.4105020202114247e-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=0.9539906103286384, f1=0.9523809523809524,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007088914280757308
step: 10, loss: 1.5377840099972673e-05

```

step: 20, loss: 1.6066996977315284e-05
step: 30, loss: 2.819957080646418e-05
step: 40, loss: 4.2559557186905295e-05
step: 50, loss: 0.00011382719094399363
step: 60, loss: 1.6666674127918668e-05
step: 70, loss: 3.0327557396958582e-05
step: 80, loss: 3.782820567721501e-05
step: 90, loss: 4.100576552446e-05
step: 100, loss: 2.886243601096794e-05
step: 110, loss: 5.9116431657457724e-05
step: 120, loss: 2.418361691525206e-05
step: 130, loss: 0.01817912980914116
step: 140, loss: 0.0011393624590709805
step: 150, loss: 1.4271439795265906e-05
step: 160, loss: 0.00024130595556925982
step: 170, loss: 0.0005429063457995653
step: 180, loss: 3.496379576972686e-05
step: 190, loss: 0.0006567402160726488
step: 200, loss: 8.270443504443392e-05
step: 210, loss: 4.496699693845585e-05
step: 220, loss: 1.931523729581386e-05
step: 230, loss: 5.991947909933515e-05
step: 240, loss: 2.4104987460304983e-05
step: 250, loss: 0.00011917008669115603
step: 260, loss: 0.001620759954676032
step: 270, loss: 1.8879516574088484e-05
step: 280, loss: 1.771346251189243e-05
step: 290, loss: 0.00046849975478835404
step: 300, loss: 5.348983177100308e-05
step: 310, loss: 5.897232404095121e-05
step: 320, loss: 0.00043154042214155197
step: 330, loss: 4.120961602893658e-05
step: 340, loss: 3.8813923310954124e-05
step: 350, loss: 2.3900609448901378e-05
step: 360, loss: 5.721872366848402e-05
step: 370, loss: 2.9516142603824846e-05
step: 380, loss: 2.612041134852916e-05
step: 390, loss: 0.026047170162200928
step: 400, loss: 0.00015871538198553026
step: 410, loss: 1.8056172848446295e-05
step: 420, loss: 5.650825914926827e-05
step: 430, loss: 2.189312181144487e-05
step: 440, loss: 9.323997073806822e-05
step: 450, loss: 3.0166111173457466e-05
step: 460, loss: 0.00022590743901673704
step: 470, loss: 2.5918674509739503e-05
step: 480, loss: 2.2287833417067304e-05
step: 490, loss: 3.278957592556253e-05

```
step: 500, loss: 1.6346393749699928e-05
step: 510, loss: 1.75458535522921e-05
step: 520, loss: 0.0003334734356030822
step: 530, loss: 3.862266385112889e-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=0.9530075187969925, f1=0.950957496496964,
best_f1=0.9458583988894032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.89389811566798e-05
step: 10, loss: 3.673839455586858e-05
step: 20, loss: 5.044449062552303e-05
step: 30, loss: 4.111746602575295e-05
step: 40, loss: 0.0952591747045517
step: 50, loss: 2.445953396090772e-05
step: 60, loss: 1.3496589417627547e-05
step: 70, loss: 4.0573115256847814e-05
step: 80, loss: 1.8059947251458652e-05
step: 90, loss: 1.8268536223331466e-05
step: 100, loss: 0.0001784664927981794
step: 110, loss: 0.0033283326774835587
step: 120, loss: 0.0005516873206943274
step: 130, loss: 0.11359520256519318
step: 140, loss: 4.042415093863383e-05
step: 150, loss: 1.9270641132607125e-05
step: 160, loss: 2.619516999402549e-05
step: 170, loss: 2.5624141926527955e-05
step: 180, loss: 1.57801405293867e-05
step: 190, loss: 2.0373323422973044e-05
step: 200, loss: 1.5098425137693994e-05
step: 210, loss: 0.0031993198208510876
step: 220, loss: 1.7210591977345757e-05
step: 230, loss: 4.5294040319276974e-05
step: 240, loss: 1.9035862351302058e-05
step: 250, loss: 2.0332208805484697e-05
step: 260, loss: 2.8395570552675053e-05
step: 270, loss: 1.7817821571952663e-05
step: 280, loss: 2.6239675207762048e-05
step: 290, loss: 1.3749913705396466e-05
step: 300, loss: 1.9266870367573574e-05
step: 310, loss: 1.831327062973287e-05
step: 320, loss: 0.0005230794777162373
step: 330, loss: 7.003036444075406e-05
step: 340, loss: 2.6881885787588544e-05
step: 350, loss: 1.4234125956136268e-05
```

```

step: 360, loss: 0.0014948582975193858
step: 370, loss: 1.9367043933016248e-05
step: 380, loss: 3.119775647064671e-05
step: 390, loss: 5.395546759245917e-05
step: 400, loss: 2.0216790289850906e-05
step: 410, loss: 2.5520863346173428e-05
step: 420, loss: 2.2958336558076553e-05
step: 430, loss: 4.7029789129737765e-05
step: 440, loss: 0.00013562748790718615
step: 450, loss: 1.4468873814621475e-05
step: 460, loss: 1.5564086425001733e-05
step: 470, loss: 0.0030431633349508047
step: 480, loss: 1.5258625353453681e-05
step: 490, loss: 1.566466198710259e-05
step: 500, loss: 3.656453191069886e-05
step: 510, loss: 1.6733776647015475e-05
step: 520, loss: 1.633521242183633e-05
step: 530, loss: 0.0005705878138542175
/usr/local/lib/python3.7/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=0.9534555712270805, f1=0.9509116409537166,
best_f1=0.9458583988894032
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 [00:28, 204.99it/s]
load_f1 = 0.9543147208121827
real_f1 = 0.9561605906783571
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.75it/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: 0.5777166485786438
step: 10, loss: 0.3767842948436737
step: 20, loss: 0.393909752368927
step: 30, loss: 0.3246627449989319
step: 40, loss: 0.16227145493030548
step: 50, loss: 0.3857443332672119
step: 60, loss: 0.20130473375320435
step: 70, loss: 0.17375335097312927
step: 80, loss: 0.21311834454536438
step: 90, loss: 0.2779470980167389
step: 100, loss: 0.3623950183391571
step: 110, loss: 0.19581739604473114
```

```

step: 120, loss: 0.16229525208473206
step: 130, loss: 0.17465505003929138
step: 140, loss: 0.15978623926639557
step: 150, loss: 0.142567977309227
step: 160, loss: 0.23383668065071106
step: 170, loss: 0.22324888408184052
step: 180, loss: 0.063225656747818
step: 190, loss: 0.21440444886684418
step: 200, loss: 0.2403349131345749
step: 210, loss: 0.23765674233436584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6878727634194831, f1=0.7054108216432866,
best_f1=0.7054108216432866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05413422733545303
step: 10, loss: 0.1322217434644699
step: 20, loss: 0.18357397615909576
step: 30, loss: 0.209498330950737
step: 40, loss: 0.1536722034215927
step: 50, loss: 0.20911003649234772
step: 60, loss: 0.40765076875686646
step: 70, loss: 0.14793042838573456
step: 80, loss: 0.1214066594839096
step: 90, loss: 0.09326708316802979
step: 100, loss: 0.008497155271470547
step: 110, loss: 0.11448734253644943
step: 120, loss: 0.1472172886133194
step: 130, loss: 0.005815309472382069
step: 140, loss: 0.191392719745636
step: 150, loss: 0.12792298197746277
step: 160, loss: 0.14006635546684265
step: 170, loss: 0.08186892420053482
step: 180, loss: 0.10994159430265427
step: 190, loss: 0.10564620047807693
step: 200, loss: 0.04420030117034912
step: 210, loss: 0.1338495910167694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.697841726618705, f1=0.7362637362637362,
best_f1=0.7362637362637362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03324390947818756
step: 10, loss: 0.18067696690559387
step: 20, loss: 0.08520819246768951
step: 30, loss: 0.2664409577846527
step: 40, loss: 0.018228385597467422
step: 50, loss: 0.0682695209980011
step: 60, loss: 0.0847587063908577
step: 70, loss: 0.03610135242342949
step: 80, loss: 0.16726244986057281
step: 90, loss: 0.023776881396770477
step: 100, loss: 0.162894144654274
step: 110, loss: 0.11367298662662506
step: 120, loss: 0.1455562263727188
step: 130, loss: 0.11242350190877914
step: 140, loss: 0.06009718403220177
step: 150, loss: 0.2354501634836197
step: 160, loss: 0.02474561147391796
step: 170, loss: 0.15079286694526672
step: 180, loss: 0.07126487046480179
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.251932293176651
step: 200, loss: 0.019497592002153397
step: 210, loss: 0.11096613854169846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7395626242544732, f1=0.7309236947791166,
best_f1=0.7309236947791166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04680875316262245
step: 10, loss: 0.02497689053416252
step: 20, loss: 0.1135302186012268
step: 30, loss: 0.048065997660160065
step: 40, loss: 0.0077456203289330006
```

```

step: 50, loss: 0.10000462830066681
step: 60, loss: 0.044293113052845
step: 70, loss: 0.17724280059337616
step: 80, loss: 0.020039528608322144
step: 90, loss: 0.05928245931863785
step: 100, loss: 0.35439029335975647
step: 110, loss: 0.09938141703605652
step: 120, loss: 0.049739085137844086
step: 130, loss: 0.16507679224014282
step: 140, loss: 0.08680977672338486
step: 150, loss: 0.12373322248458862
step: 160, loss: 0.01576974056661129
step: 170, loss: 0.19796684384346008
step: 180, loss: 0.16448011994361877
step: 190, loss: 0.05516427010297775
step: 200, loss: 0.05477454885840416
step: 210, loss: 0.14300857484340668
/usr/local/lib/python3.7/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=0.7252747252747251, f1=0.7248576850094877,
best_f1=0.7309236947791166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11009740829467773
step: 10, loss: 0.02801435999572277
step: 20, loss: 0.014818824827671051
step: 30, loss: 0.015273923985660076
step: 40, loss: 0.007282511796802282
step: 50, loss: 0.10304022580385208
step: 60, loss: 0.04908955469727516
step: 70, loss: 0.0023289918899536133
step: 80, loss: 0.012996991164982319
step: 90, loss: 0.07051289081573486
step: 100, loss: 0.004201819188892841
step: 110, loss: 0.15980495512485504
step: 120, loss: 0.12418852001428604
step: 130, loss: 0.05339444428682327
step: 140, loss: 0.016961926594376564
step: 150, loss: 0.03724488988518715
step: 160, loss: 0.13609158992767334
step: 170, loss: 0.05193169414997101
step: 180, loss: 0.1244221180677414
step: 190, loss: 0.014050749130547047
step: 200, loss: 0.07997381687164307
step: 210, loss: 0.01996145024895668
/usr/local/lib/python3.7/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=0.7225325884543762, f1=0.7230769230769231,
best_f1=0.7309236947791166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03375571593642235
step: 10, loss: 0.03548748046159744
step: 20, loss: 0.0030318512581288815
step: 30, loss: 0.0018283212557435036
step: 40, loss: 0.02540784887969494
step: 50, loss: 0.00848180241882801
step: 60, loss: 0.07977651059627533
step: 70, loss: 0.04592316970229149
step: 80, loss: 0.0804705023765564
step: 90, loss: 0.10143356770277023
step: 100, loss: 0.0013809714000672102
step: 110, loss: 0.037679728120565414
step: 120, loss: 0.1281725913286209
step: 130, loss: 0.04030416160821915
step: 140, loss: 0.08883819729089737
step: 150, loss: 0.008739441633224487
step: 160, loss: 0.0030974240507930517
step: 170, loss: 0.04333390295505524
step: 180, loss: 0.01639191247522831
step: 190, loss: 0.06977298855781555
step: 200, loss: 0.0015666995896026492
step: 210, loss: 0.03535314276814461
/usr/local/lib/python3.7/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=0.7217741935483871, f1=0.7145790554414784,
best_f1=0.7309236947791166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15262536704540253
step: 10, loss: 0.022942475974559784
step: 20, loss: 0.006845146883279085
step: 30, loss: 0.007732827216386795
step: 40, loss: 0.009191370569169521
step: 50, loss: 0.04963986575603485
step: 60, loss: 0.048652395606040955
step: 70, loss: 0.00986085832118988
step: 80, loss: 0.03712482005357742
step: 90, loss: 0.0013331001391634345
step: 100, loss: 0.0005211058305576444

```

```

step: 110, loss: 0.13272225856781006
step: 120, loss: 0.07510627061128616
step: 130, loss: 0.04077071696519852
step: 140, loss: 0.020871473476290703
step: 150, loss: 0.01720534637570381
step: 160, loss: 0.031703319400548935
step: 170, loss: 0.002974832197651267
step: 180, loss: 0.017763681709766388
step: 190, loss: 0.07516967505216599
step: 200, loss: 0.0011310568079352379
step: 210, loss: 0.0034292489290237427
/usr/local/lib/python3.7/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=0.71919191919192, f1=0.7131147540983607,
best_f1=0.7309236947791166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04030522704124451
step: 10, loss: 0.054423462599515915
step: 20, loss: 0.0025380169972777367
step: 30, loss: 0.0706789642572403
step: 40, loss: 0.0020821725483983755
step: 50, loss: 0.004194342531263828
step: 60, loss: 0.10332103818655014
step: 70, loss: 0.0555107481777668
step: 80, loss: 0.014669411815702915
step: 90, loss: 0.0015497267013415694
step: 100, loss: 0.010447666980326176
step: 110, loss: 0.12928280234336853
step: 120, loss: 0.0004238089604768902
step: 130, loss: 0.0007169494056142867
step: 140, loss: 0.035229794681072235
step: 150, loss: 0.004574637394398451
step: 160, loss: 0.03549599274992943
step: 170, loss: 0.01138587761670351
step: 180, loss: 0.08994752913713455
step: 190, loss: 0.12396083772182465
step: 200, loss: 0.0028479984030127525
step: 210, loss: 0.4894970655441284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7431906614785992, f1=0.7233201581027668,
best_f1=0.7233201581027668
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002196140121668577
step: 10, loss: 0.008651236072182655
step: 20, loss: 0.01089970301836729
step: 30, loss: 0.0027056357357650995
step: 40, loss: 0.055736299604177475
step: 50, loss: 0.0612761452794075
step: 60, loss: 0.07070315629243851
step: 70, loss: 0.014130156487226486
step: 80, loss: 0.0036809351295232773
step: 90, loss: 0.01965135894715786
step: 100, loss: 0.001703142188489437
step: 110, loss: 0.025761054828763008
step: 120, loss: 0.003384545911103487
step: 130, loss: 0.005504367873072624
step: 140, loss: 0.027316853404045105
step: 150, loss: 0.14952623844146729
step: 160, loss: 0.0005385951953940094
step: 170, loss: 0.0015222678193822503
step: 180, loss: 0.05943623557686806
step: 190, loss: 0.007940874435007572
step: 200, loss: 0.06265593320131302
step: 210, loss: 0.005105436313897371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.746987951807229, f1=0.7222222222222222,
best_f1=0.7222222222222222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040214646607637405
step: 10, loss: 0.016213366761803627
step: 20, loss: 0.0012008455814793706
step: 30, loss: 0.058726247400045395
step: 40, loss: 0.0002886336005758494
step: 50, loss: 0.001523183542303741
step: 60, loss: 0.03263525292277336
step: 70, loss: 0.0621015690267086
step: 80, loss: 0.015344260260462761

```



```

step: 90, loss: 0.03085022047162056
step: 100, loss: 0.0049106664955616
step: 110, loss: 0.04408751428127289
step: 120, loss: 0.0005241595790721476
step: 130, loss: 0.016360817477107048
step: 140, loss: 0.009366406127810478
step: 150, loss: 0.012616278603672981
step: 160, loss: 0.0026736892759799957
step: 170, loss: 0.0005444595590233803
step: 180, loss: 0.024175744503736496
step: 190, loss: 0.0688055157661438
step: 200, loss: 0.006620810832828283
step: 210, loss: 0.045558035373687744
/usr/local/lib/python3.7/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=0.733606557377049, f1=0.7128309572301426,
best_f1=0.7222222222222222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03153665363788605
step: 10, loss: 0.018055012449622154
step: 20, loss: 0.006081497296690941
step: 30, loss: 0.00020907340513076633
step: 40, loss: 0.039792466908693314
step: 50, loss: 0.04000261053442955
step: 60, loss: 0.023880651220679283
step: 70, loss: 0.010106606408953667
step: 80, loss: 0.021501537412405014
step: 90, loss: 0.01851872354745865
step: 100, loss: 0.003902625758200884
step: 110, loss: 0.142234206199646
step: 120, loss: 0.06359609216451645
step: 130, loss: 0.0023543040733784437
step: 140, loss: 0.00014553497021552175
step: 150, loss: 0.006622124928981066
step: 160, loss: 0.0013707849429920316
step: 170, loss: 0.00395914725959301
step: 180, loss: 0.0003609596751630306
step: 190, loss: 0.0015090532833710313
step: 200, loss: 0.0002815455663949251
step: 210, loss: 0.00015199645713437349
/usr/local/lib/python3.7/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=0.7280334728033474, f1=0.727659574468085,
best_f1=0.7222222222222222

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029556220397353172
step: 10, loss: 0.00021014909725636244
step: 20, loss: 0.02540973201394081
step: 30, loss: 0.01650545932352543
step: 40, loss: 0.0013069723499938846
step: 50, loss: 0.002822508569806814
step: 60, loss: 0.0020479599479585886
step: 70, loss: 0.0009821610292419791
step: 80, loss: 0.0008802515221759677
step: 90, loss: 0.02510547824203968
step: 100, loss: 0.004012836143374443
step: 110, loss: 0.0009610528941266239
step: 120, loss: 0.0013190667377784848
step: 130, loss: 0.0005164356552995741
step: 140, loss: 0.0004074382013641298
step: 150, loss: 0.010537682101130486
step: 160, loss: 0.0006191795691847801
step: 170, loss: 0.006599028594791889
step: 180, loss: 0.08678825199604034
step: 190, loss: 0.015061340294778347
step: 200, loss: 0.04014592990279198
step: 210, loss: 0.016624081879854202
/usr/local/lib/python3.7/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=0.733050847457627, f1=0.7112068965517242,
best_f1=0.7222222222222222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03304314985871315
step: 10, loss: 0.0007953916792757809
step: 20, loss: 0.041964948177337646
step: 30, loss: 0.02717147022485733
step: 40, loss: 0.026379257440567017
step: 50, loss: 0.002448741113767028
step: 60, loss: 0.00033711487776599824
step: 70, loss: 0.04961172118782997
step: 80, loss: 0.011162267066538334
step: 90, loss: 0.00015159918984863907
step: 100, loss: 0.0008663117769174278
step: 110, loss: 0.0001610970648471266
step: 120, loss: 0.0002777522895485163
step: 130, loss: 0.00024152548576239496
step: 140, loss: 0.03204052522778511

```

```

step: 150, loss: 0.0007458659820258617
step: 160, loss: 0.015719085931777954
step: 170, loss: 0.000152006366988644
step: 180, loss: 0.044675063341856
step: 190, loss: 0.0008271902333945036
step: 200, loss: 0.000199384186998941
step: 210, loss: 0.08024206757545471
/usr/local/lib/python3.7/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=0.7346938775510206, f1=0.7080745341614907,
best_f1=0.7222222222222222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.569935405626893e-05
step: 10, loss: 0.03110901080071926
step: 20, loss: 0.0010193450143560767
step: 30, loss: 0.00015841078129597008
step: 40, loss: 0.00031219387892633677
step: 50, loss: 0.02675236016511917
step: 60, loss: 0.0008727338863536716
step: 70, loss: 0.00015187007375061512
step: 80, loss: 0.00016371933452319354
step: 90, loss: 0.0011382558150216937
step: 100, loss: 0.030658695846796036
step: 110, loss: 0.0006131260306574404
step: 120, loss: 0.046223439276218414
step: 130, loss: 0.0007412094855681062
step: 140, loss: 0.02803356945514679
step: 150, loss: 0.0005768979317508638
step: 160, loss: 0.0013988196151331067
step: 170, loss: 0.0006216774927452207
step: 180, loss: 0.00041154026985168457
step: 190, loss: 0.0006037629209458828
step: 200, loss: 0.00014827511040493846
step: 210, loss: 0.00480555510148406
/usr/local/lib/python3.7/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=0.7292110874200425, f1=0.7068965517241379,
best_f1=0.7222222222222222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001571361062815413
step: 10, loss: 6.218083581188694e-05
step: 20, loss: 0.00024560175370424986

```

```

step: 30, loss: 0.025355223566293716
step: 40, loss: 0.003932857420295477
step: 50, loss: 0.00015576356963720173
step: 60, loss: 0.000304901011986658
step: 70, loss: 0.001771290204487741
step: 80, loss: 0.00016838214651215822
step: 90, loss: 0.019335834309458733
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.00024253176525235176
step: 110, loss: 0.0003399980196263641
step: 120, loss: 0.00022069655824452639
step: 130, loss: 0.00023947075533214957
step: 140, loss: 8.614722173660994e-05
step: 150, loss: 0.00010658930841600522
step: 160, loss: 0.00037425453774631023
step: 170, loss: 0.0027221208438277245
step: 180, loss: 0.0002046285808319226
step: 190, loss: 0.0001584790152264759
step: 200, loss: 0.0035240305587649345
step: 210, loss: 0.0008479764801450074
/usr/local/lib/python3.7/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=0.7295597484276728, f1=0.7031578947368422,
best_f1=0.7222222222222222
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, 285.74it/s]
load_f1 = 0.7384615384615384
real_f1 = 0.7384615384615384
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, 194.11it/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/
```

```
Downloading: 100% 433/433 [00:00<00:00, 368kB/s]  
Downloading: 100% 232k/232k [00:00<00:00, 256kB/s]  
Downloading: 100% 440M/440M [00:06<00:00, 67.3MB/s]  
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32  
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

```
step: 0, loss: 0.5777714252471924  
step: 10, loss: 0.36897197365760803  
step: 20, loss: 0.28100502490997314  
step: 30, loss: 0.4311293959617615  
step: 40, loss: 0.42962518334388733
```

```

step: 50, loss: 0.31618162989616394
step: 60, loss: 0.25001218914985657
step: 70, loss: 0.24974225461483002
step: 80, loss: 0.21498481929302216
step: 90, loss: 0.24167747795581818
step: 100, loss: 0.282459020614624
step: 110, loss: 0.43239396810531616
step: 120, loss: 0.16525304317474365
step: 130, loss: 0.14911669492721558
step: 140, loss: 0.04196993634104729
step: 150, loss: 0.12906421720981598
step: 160, loss: 0.06446806341409683
step: 170, loss: 0.1200663149356842
step: 180, loss: 0.03385131061077118
step: 190, loss: 0.13549794256687164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7272727272727272, f1=0.7349081364829396,
best_f1=0.7349081364829396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2126997709274292
step: 10, loss: 0.04676397144794464
step: 20, loss: 0.05387983098626137
step: 30, loss: 0.21893487870693207
step: 40, loss: 0.185978963971138
step: 50, loss: 0.1375555843114853
step: 60, loss: 0.30569595098495483
step: 70, loss: 0.07350996881723404
step: 80, loss: 0.12005005031824112
step: 90, loss: 0.11653117835521698
step: 100, loss: 0.02805051952600479
step: 110, loss: 0.1456328183412552
step: 120, loss: 0.20355211198329926
step: 130, loss: 0.06346728652715683
step: 140, loss: 0.05077280476689339
step: 150, loss: 0.0437081977725029
step: 160, loss: 0.025644246488809586
step: 170, loss: 0.06985678523778915
step: 180, loss: 0.07434138655662537
step: 190, loss: 0.027573052793741226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7904509283819628, f1=0.7760416666666666,
best_f1=0.7760416666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022380398586392403
step: 10, loss: 0.051625847816467285
step: 20, loss: 0.01517017837613821
step: 30, loss: 0.00772339990362525
step: 40, loss: 0.012953064404428005
step: 50, loss: 0.06809066236019135
step: 60, loss: 0.008399847894906998
step: 70, loss: 0.0970347598195076
step: 80, loss: 0.06854669004678726
step: 90, loss: 0.04580335691571236
step: 100, loss: 0.005640239454805851
step: 110, loss: 0.0020114679355174303
step: 120, loss: 0.019282266497612
step: 130, loss: 0.006206152029335499
step: 140, loss: 0.03328137844800949
step: 150, loss: 0.071266770362854
step: 160, loss: 0.03787287324666977
step: 170, loss: 0.09676948189735413
step: 180, loss: 0.03244457393884659
step: 190, loss: 0.035042330622673035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8119891008174387, f1=0.8228882833787465,
best_f1=0.8228882833787465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012227733619511127
step: 10, loss: 0.11524228006601334
step: 20, loss: 0.03921617195010185
step: 30, loss: 0.004952843766659498
step: 40, loss: 0.08516990393400192

```



```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.19022582471370697
step: 60, loss: 0.005878178868442774
step: 70, loss: 0.1044929027557373
step: 80, loss: 0.15312696993350983
step: 90, loss: 0.022838452830910683
step: 100, loss: 0.08579698950052261
step: 110, loss: 0.008170763961970806
step: 120, loss: 0.008109680376946926
step: 130, loss: 0.03736163303256035
step: 140, loss: 0.004179120063781738
step: 150, loss: 0.009138678200542927
step: 160, loss: 0.0015398397808894515
step: 170, loss: 0.01988399401307106
step: 180, loss: 0.01163721363991499
step: 190, loss: 0.03590608388185501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8125, f1=0.8219895287958114, best_f1=0.8219895287958114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041313908994197845
step: 10, loss: 0.0046470589004457
step: 20, loss: 0.007546477019786835
step: 30, loss: 0.02254493534564972
step: 40, loss: 0.04122130200266838
step: 50, loss: 0.16475461423397064
step: 60, loss: 0.08744761347770691
step: 70, loss: 0.002913103671744466
step: 80, loss: 0.0016220106044784188
step: 90, loss: 0.27174898982048035
step: 100, loss: 0.0009420174756087363
step: 110, loss: 0.005492581054568291
step: 120, loss: 0.0021564848721027374
step: 130, loss: 0.22573523223400116
step: 140, loss: 0.0023926973808556795
step: 150, loss: 0.018533505499362946
step: 160, loss: 0.015899771824479103
step: 170, loss: 0.002211957471445203
step: 180, loss: 0.03367488086223602
step: 190, loss: 0.012502101249992847
/usr/local/lib/python3.7/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=0.7948051948051948, f1=0.8, best_f1=0.8219895287958114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019631264731287956
step: 10, loss: 0.04390735179185867
step: 20, loss: 0.00639914209023118
step: 30, loss: 0.06677138805389404
step: 40, loss: 0.0953676775097847
step: 50, loss: 0.0034020300954580307
step: 60, loss: 0.017257103696465492
step: 70, loss: 0.012605485506355762
step: 80, loss: 0.0649326741695404
step: 90, loss: 0.0008068372262641788
step: 100, loss: 0.0004603216366376728
step: 110, loss: 0.002668239176273346
step: 120, loss: 0.0025435490533709526
step: 130, loss: 0.019222622737288475
step: 140, loss: 0.0022031923290342093
step: 150, loss: 0.0004947854322381318
step: 160, loss: 0.0018226299434900284
step: 170, loss: 0.005709170829504728
step: 180, loss: 0.04365646839141846
step: 190, loss: 0.019726702943444252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8142493638676844, f1=0.8163265306122449,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004944984335452318
step: 10, loss: 0.010238208808004856
step: 20, loss: 0.011431172490119934
step: 30, loss: 0.0024664271622896194
step: 40, loss: 0.002332910895347595
step: 50, loss: 0.008537957444787025
step: 60, loss: 0.00235443701967597
step: 70, loss: 0.015079999342560768
step: 80, loss: 0.018280811607837677
step: 90, loss: 0.0009370011393912137

```

```

step: 100, loss: 0.0005030978936702013
step: 110, loss: 0.02239266224205494
step: 120, loss: 0.0008976557874120772
step: 130, loss: 0.0006408469052985311
step: 140, loss: 0.0011366402031853795
step: 150, loss: 0.007864933460950851
step: 160, loss: 0.016520844772458076
step: 170, loss: 0.0016028976533561945
step: 180, loss: 0.0005794455064460635
step: 190, loss: 0.0013610547175630927
/usr/local/lib/python3.7/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=0.8043478260869565, f1=0.7823691460055096,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01741882972419262
step: 10, loss: 0.020843718200922012
step: 20, loss: 0.00046938180457800627
step: 30, loss: 0.0008603085298091173
step: 40, loss: 0.006163395941257477
step: 50, loss: 0.0010876859305426478
step: 60, loss: 0.0004221523704472929
step: 70, loss: 0.006631829775869846
step: 80, loss: 0.00031068548560142517
step: 90, loss: 0.0003754344361368567
step: 100, loss: 0.04354109987616539
step: 110, loss: 0.20030002295970917
step: 120, loss: 0.0008169515058398247
step: 130, loss: 0.002302302746102214
step: 140, loss: 0.0007480430649593472
step: 150, loss: 0.005290776491165161
step: 160, loss: 0.0004917976912111044
step: 170, loss: 0.0005366828991100192
step: 180, loss: 0.00240463949739933
step: 190, loss: 0.000673739006742835
/usr/local/lib/python3.7/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=0.8109452736318408, f1=0.7938931297709924,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006675793556496501
step: 10, loss: 0.04802296310663223

```

```

step: 20, loss: 0.0012464472092688084
step: 30, loss: 0.000756678229663521
step: 40, loss: 0.00043973029823973775
step: 50, loss: 0.0021328190341591835
step: 60, loss: 0.0032388202380388975
step: 70, loss: 0.003773168660700321
step: 80, loss: 0.0006928267539478838
step: 90, loss: 0.00119685847312212
step: 100, loss: 0.04867299273610115
step: 110, loss: 0.0007594658527523279
step: 120, loss: 0.002002730965614319
step: 130, loss: 0.057748693972826004
step: 140, loss: 0.0015069048386067152
step: 150, loss: 0.0008006257121451199
step: 160, loss: 0.00020178455451969057
step: 170, loss: 0.001204690895974636
step: 180, loss: 0.027471594512462616
step: 190, loss: 0.0014301618793979287
/usr/local/lib/python3.7/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=0.8, f1=0.8071065989847714, best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046141058555804193
step: 10, loss: 0.0004409262619446963
step: 20, loss: 0.006230646278709173
step: 30, loss: 0.0001391273835906759
step: 40, loss: 0.0005454698693938553
step: 50, loss: 0.0004676641256082803
step: 60, loss: 0.0011127474717795849
step: 70, loss: 0.04250446707010269
step: 80, loss: 0.00012715595948975533
step: 90, loss: 0.00012552272528409958
step: 100, loss: 0.0003688986471388489
step: 110, loss: 0.0015222131041809916
step: 120, loss: 0.11480609327554703
step: 130, loss: 0.00020083518757019192
step: 140, loss: 0.0011054161004722118
step: 150, loss: 0.00036368719884194434
step: 160, loss: 0.0022493829019367695
step: 170, loss: 0.00016776527627371252
step: 180, loss: 0.00020702145411632955
step: 190, loss: 0.0001816914154915139
/usr/local/lib/python3.7/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=0.8021680216802168, f1=0.8324324324324325,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013137716450728476
step: 10, loss: 0.0002123151789419353
step: 20, loss: 0.002691042609512806
step: 30, loss: 0.03488251194357872
step: 40, loss: 0.0001282214216189459
step: 50, loss: 0.00042918979306705296
step: 60, loss: 0.00015198258915916085
step: 70, loss: 0.000990497530438006
step: 80, loss: 0.00033544510370120406
step: 90, loss: 0.00012390992196742445
step: 100, loss: 0.00015212352445814759
step: 110, loss: 0.00024511138326488435
step: 120, loss: 0.000992042594589293
step: 130, loss: 0.00019575589976739138
step: 140, loss: 0.0009970094542950392
step: 150, loss: 0.00012522184988483787
step: 160, loss: 0.0003358208923600614
step: 170, loss: 0.00030878433608449996
step: 180, loss: 0.0012350097531452775
step: 190, loss: 0.0003691067104227841
/usr/local/lib/python3.7/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=0.8009708737864077, f1=0.8112244897959184,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013118362403474748
step: 10, loss: 0.004684319254010916
step: 20, loss: 0.0003255657502450049
step: 30, loss: 0.00028184056282043457
step: 40, loss: 0.0009990582475438714
step: 50, loss: 0.0009602111531421542
step: 60, loss: 0.00037585728568956256
step: 70, loss: 0.00011441379319876432
step: 80, loss: 0.00023147212050389498
step: 90, loss: 0.00012814342335332185
step: 100, loss: 0.0025570651050657034
step: 110, loss: 0.00034344205050729215
step: 120, loss: 0.00017148868937510997
step: 130, loss: 0.00016386673087254167
step: 140, loss: 0.00044610220356844366

```

```

step: 150, loss: 0.0023294698912650347
step: 160, loss: 0.00013668990868609399
step: 170, loss: 0.0010340535081923008
step: 180, loss: 8.623849862487987e-05
step: 190, loss: 0.0002191897656302899
/usr/local/lib/python3.7/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=0.8087167070217917, f1=0.8151898734177214,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000373867922462523
step: 10, loss: 0.00010744817700469866
step: 20, loss: 0.00023064254492055625
step: 30, loss: 0.00024578924058005214
step: 40, loss: 0.00011204845941392705
step: 50, loss: 0.00029412275762297213
step: 60, loss: 0.00015230207645799965
step: 70, loss: 0.0023284442722797394
step: 80, loss: 0.00027317178319208324
step: 90, loss: 0.0002058020036201924
step: 100, loss: 0.0004120134108234197
step: 110, loss: 0.00011981371062574908
step: 120, loss: 0.0001888517290353775
step: 130, loss: 0.0001357102009933442
step: 140, loss: 0.0015988191589713097
step: 150, loss: 8.488543971907347e-05
step: 160, loss: 0.00015196156164165586
step: 170, loss: 0.00019371228700038046
step: 180, loss: 0.0004002869827672839
step: 190, loss: 0.0011912673944607377
/usr/local/lib/python3.7/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=0.8041237113402062, f1=0.8063660477453581,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017199036665260792
step: 10, loss: 0.00018210598500445485
step: 20, loss: 0.00030358764342963696
step: 30, loss: 8.670528040966019e-05
step: 40, loss: 0.0015189102850854397
step: 50, loss: 0.00014324374205898494
step: 60, loss: 0.0004238304099999368

```

```

step: 70, loss: 0.00016280192357953638
step: 80, loss: 0.00025701409322209656
step: 90, loss: 0.0001318172726314515
step: 100, loss: 0.001091143349185586
step: 110, loss: 0.00023205451725516468
step: 120, loss: 0.0002394929324509576
step: 130, loss: 0.00010607328295009211
step: 140, loss: 0.00019294294179417193
step: 150, loss: 0.0009978958405554295
step: 160, loss: 0.00012919232540298253
step: 170, loss: 0.00025218870723620057
step: 180, loss: 0.00012896143016405404
step: 190, loss: 0.00015914418327156454
/usr/local/lib/python3.7/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=0.8051948051948051, f1=0.8201058201058201,
best_f1=0.8163265306122449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032845553942024708
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 5.9911541029578075e-05
step: 20, loss: 8.564884774386883e-05
step: 30, loss: 8.318630716530606e-05
step: 40, loss: 0.00010311192454537377
step: 50, loss: 0.00011463023110991344
step: 60, loss: 0.06588377803564072
step: 70, loss: 0.010934503749012947
step: 80, loss: 0.000996735063381493
step: 90, loss: 0.0003251066373195499
step: 100, loss: 0.00015509003424085677
step: 110, loss: 0.00020816686446778476
step: 120, loss: 9.262029925594106e-05
step: 130, loss: 0.0004575576458591968
step: 140, loss: 0.0004014630103483796
step: 150, loss: 0.0004146405844949186
step: 160, loss: 0.00031773876980878413
step: 170, loss: 0.19474159181118011
step: 180, loss: 0.0004781207535415888
step: 190, loss: 0.00020240229787304997
/usr/local/lib/python3.7/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=0.8051948051948051, f1=0.8148148148148148,
best_f1=0.8163265306122449
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 [00:11, 176.06it/s]
load_f1 = 0.6208791208791209
real_f1 = 0.6049046321525886
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, 192.93it/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: 0.6337838172912598
step: 10, loss: 0.36209648847579956
step: 20, loss: 0.2985975444316864
step: 30, loss: 0.38312557339668274
step: 40, loss: 0.2808328866958618
step: 50, loss: 0.24997544288635254
step: 60, loss: 0.2965647280216217
step: 70, loss: 0.366421103477478
step: 80, loss: 0.31820032000541687
step: 90, loss: 0.1460198312997818
step: 100, loss: 0.13393673300743103
step: 110, loss: 0.2310798317193985
step: 120, loss: 0.19617100059986115
step: 130, loss: 0.05823595076799393
step: 140, loss: 0.12130444496870041
step: 150, loss: 0.2851525843143463
step: 160, loss: 0.1079162061214447
step: 170, loss: 0.26727786660194397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7777777777777778, f1=0.730310262529833,
best_f1=0.730310262529833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06952472776174545
step: 10, loss: 0.20349016785621643

```

```

step: 20, loss: 0.07050527632236481
step: 30, loss: 0.20065069198608398
step: 40, loss: 0.10320912301540375
step: 50, loss: 0.06580093502998352
step: 60, loss: 0.1359851360321045
step: 70, loss: 0.06952585279941559
step: 80, loss: 0.11944618076086044
step: 90, loss: 0.0952942967414856
step: 100, loss: 0.08009131997823715
step: 110, loss: 0.04024225100874901
step: 120, loss: 0.02044355869293213
step: 130, loss: 0.15782058238983154
step: 140, loss: 0.20167559385299683
step: 150, loss: 0.19806943833827972
step: 160, loss: 0.09406087547540665
step: 170, loss: 0.06504655629396439
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8165938864628821, f1=0.7899159663865547,
best_f1=0.7899159663865547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08707845956087112
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.08713193237781525
step: 20, loss: 0.05590403079986572
step: 30, loss: 0.13343173265457153
step: 40, loss: 0.13292653858661652
step: 50, loss: 0.08540035039186478
step: 60, loss: 0.07345298677682877
step: 70, loss: 0.04981683939695358
step: 80, loss: 0.040596477687358856
step: 90, loss: 0.08711156249046326
step: 100, loss: 0.01428010780364275
step: 110, loss: 0.06230253726243973
step: 120, loss: 0.1271178424358368
step: 130, loss: 0.031141960993409157
step: 140, loss: 0.14455175399780273
step: 150, loss: 0.025234706699848175
step: 160, loss: 0.04076508805155754
step: 170, loss: 0.09432333707809448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8217821782178217, f1=0.831353919239905,
best_f1=0.831353919239905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02066061832010746
step: 10, loss: 0.011866189539432526
step: 20, loss: 0.03258317708969116
step: 30, loss: 0.023756342008709908
step: 40, loss: 0.0018477411940693855
step: 50, loss: 0.05389060825109482
step: 60, loss: 0.03421078622341156
step: 70, loss: 0.0016130984295159578
step: 80, loss: 0.033679522573947906
step: 90, loss: 0.0577206090092659
step: 100, loss: 0.1134243980050087
step: 110, loss: 0.05988847091794014
step: 120, loss: 0.00647288653999567
step: 130, loss: 0.041308801621198654
step: 140, loss: 0.051407620310783386
step: 150, loss: 0.012467978522181511
step: 160, loss: 0.0446760468184948
step: 170, loss: 0.0017173605738207698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8564356435643565, f1=0.8341232227488151,
best_f1=0.8341232227488151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004094733856618404
step: 10, loss: 0.006488861050456762
step: 20, loss: 0.014501444064080715
step: 30, loss: 0.07518712431192398
step: 40, loss: 0.010853762738406658
step: 50, loss: 0.029370887205004692
step: 60, loss: 0.004200090654194355

```

```

step: 70, loss: 0.034554727375507355
step: 80, loss: 0.05678737163543701
step: 90, loss: 0.10400427877902985
step: 100, loss: 0.04840855672955513
step: 110, loss: 0.026979943737387657
step: 120, loss: 0.004062823485583067
step: 130, loss: 0.005803290754556656
step: 140, loss: 0.016843847930431366
step: 150, loss: 0.010489373467862606
step: 160, loss: 0.017719479277729988
step: 170, loss: 0.012035616673529148
/usr/local/lib/python3.7/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=0.8541666666666666, f1=0.8592592592592592,
best_f1=0.8341232227488151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000898175232578069
step: 10, loss: 0.007405315991491079
step: 20, loss: 0.0030935544054955244
step: 30, loss: 0.0007656080997548997
step: 40, loss: 0.001440506661310792
step: 50, loss: 0.06488525867462158
step: 60, loss: 0.04110216349363327
step: 70, loss: 0.047272324562072754
step: 80, loss: 0.058794934302568436
step: 90, loss: 0.02507549151778221
step: 100, loss: 0.00500823138281703
step: 110, loss: 0.0017551702912896872
step: 120, loss: 0.00875056628137827
step: 130, loss: 0.04381589964032173
step: 140, loss: 0.014069762080907822
step: 150, loss: 0.011058464646339417
step: 160, loss: 0.043027009814977646
step: 170, loss: 0.004612009506672621
/usr/local/lib/python3.7/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=0.847926267281106, f1=0.8061674008810573,
best_f1=0.8341232227488151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024005421437323093
step: 10, loss: 0.010499552823603153
step: 20, loss: 0.01075675804167986

```

```

step: 30, loss: 0.0117802107706666
step: 40, loss: 0.0023842654190957546
step: 50, loss: 0.044192783534526825
step: 60, loss: 0.003563366597518325
step: 70, loss: 0.0008111105998978019
step: 80, loss: 0.014153048396110535
step: 90, loss: 0.0003351642226334661
step: 100, loss: 0.002912439638748765
step: 110, loss: 0.02226010337471962
step: 120, loss: 0.010313729755580425
step: 130, loss: 0.17563678324222565
step: 140, loss: 0.0008711732807569206
step: 150, loss: 0.001120410393923521
step: 160, loss: 0.0002595153637230396
step: 170, loss: 0.02601773850619793
/usr/local/lib/python3.7/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=0.8505154639175257, f1=0.8457711442786069,
best_f1=0.8341232227488151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019320950377732515
step: 10, loss: 0.000332870171405375
step: 20, loss: 0.00020780257182195783
step: 30, loss: 0.031979355961084366
step: 40, loss: 0.0002000113163376227
step: 50, loss: 0.0007462847279384732
step: 60, loss: 0.0011076544178649783
step: 70, loss: 0.0023284917697310448
step: 80, loss: 0.009665535762906075
step: 90, loss: 0.003953250590711832
step: 100, loss: 0.018172096461057663
step: 110, loss: 0.08304569125175476
step: 120, loss: 0.0009271460003219545
step: 130, loss: 0.0003567440726328641
step: 140, loss: 0.0012487967032939196
step: 150, loss: 0.013158975169062614
step: 160, loss: 0.03886732459068298
step: 170, loss: 0.02260412834584713
/usr/local/lib/python3.7/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=0.8442211055276382, f1=0.85012285012285,
best_f1=0.8341232227488151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026150750927627087
step: 10, loss: 0.005072159692645073
step: 20, loss: 0.0008386397385038435
step: 30, loss: 0.008325884118676186
step: 40, loss: 0.0006166203529573977
step: 50, loss: 0.00018277890922036022
step: 60, loss: 0.0025448028463870287
step: 70, loss: 0.03567865490913391
step: 80, loss: 0.00027291415608488023
step: 90, loss: 0.017226392403244972
step: 100, loss: 0.01640121266245842
step: 110, loss: 0.0003432748490013182
step: 120, loss: 0.007374962326139212
step: 130, loss: 0.0287010595202446
step: 140, loss: 0.0003083658230025321
step: 150, loss: 0.00028745713643729687
step: 160, loss: 0.027171004563570023
step: 170, loss: 0.026080654934048653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8599508599508598, f1=0.8530805687203792,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09976749867200851
step: 10, loss: 0.0002274468424730003
step: 20, loss: 0.0599852129817009
step: 30, loss: 0.005430157762020826
step: 40, loss: 0.0002200261369580403
step: 50, loss: 0.08174319565296173
step: 60, loss: 0.00010296914115315303
step: 70, loss: 0.008200136944651604
step: 80, loss: 0.009560422040522099
step: 90, loss: 0.0056471130810678005
step: 100, loss: 0.00023818755289539695
step: 110, loss: 0.0009319098317064345
step: 120, loss: 0.0002195286942878738
step: 130, loss: 0.00032970504253171384
step: 140, loss: 0.0006403623847290874
step: 150, loss: 0.00024714510072954
step: 160, loss: 0.00429625203832984

```

```

step: 170, loss: 0.0012340112589299679
/usr/local/lib/python3.7/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=0.8523809523809525, f1=0.8294930875576036,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012012735940515995
step: 10, loss: 0.0037601797375828028
step: 20, loss: 0.000133450172143057
step: 30, loss: 0.00011285279470030218
step: 40, loss: 0.00013902092177886516
step: 50, loss: 0.0029267265927046537
step: 60, loss: 0.00807107426226139
step: 70, loss: 0.00015281020023394376
step: 80, loss: 0.00016637492808513343
step: 90, loss: 0.00014362837828230113
step: 100, loss: 9.940918971551582e-05
step: 110, loss: 0.0056236665695905685
step: 120, loss: 0.0002634781994856894
step: 130, loss: 0.0005473757046274841
step: 140, loss: 0.004098293371498585
step: 150, loss: 0.0011546246241778135
step: 160, loss: 0.004109250381588936
step: 170, loss: 0.01593845896422863
/usr/local/lib/python3.7/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=0.8496420047732696, f1=0.8440366972477065,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03259364515542984
step: 10, loss: 0.0006628151750192046
step: 20, loss: 0.00014041532995179296
step: 30, loss: 0.00016757624689489603
step: 40, loss: 0.00021236306929495186
step: 50, loss: 0.00018870545318350196
step: 60, loss: 0.00023034418700262904
step: 70, loss: 0.08684052526950836
step: 80, loss: 0.00017246484640054405
step: 90, loss: 0.00034240196691825986
step: 100, loss: 0.00211469829082489
step: 110, loss: 0.00020039269293192774
step: 120, loss: 0.01450982503592968

```



```

step: 130, loss: 0.01007701363414526
step: 140, loss: 0.009579028934240341
step: 150, loss: 0.003901419695466757
step: 160, loss: 0.00020426178525667638
step: 170, loss: 0.0002441935648676008
/usr/local/lib/python3.7/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=0.8415841584158416, f1=0.8467153284671532,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015281407162547112
step: 10, loss: 0.0008269631071016192
step: 20, loss: 0.0001530369045212865
step: 30, loss: 0.00017017990467138588
step: 40, loss: 0.0008064459543675184
step: 50, loss: 0.00019479732145555317
step: 60, loss: 0.004606155678629875
step: 70, loss: 0.0003203728701919317
step: 80, loss: 0.0008639793377369642
step: 90, loss: 0.00018464625463820994
step: 100, loss: 0.00025157121126540005
step: 110, loss: 0.00016183205298148096
step: 120, loss: 0.032206375151872635
step: 130, loss: 0.0015121730975806713
step: 140, loss: 0.00019135061302222311
step: 150, loss: 0.0009303954429924488
step: 160, loss: 0.00033555389381945133
step: 170, loss: 0.0013200825778767467
/usr/local/lib/python3.7/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=0.8312342569269522, f1=0.8467153284671532,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002650693932082504
step: 10, loss: 0.00010935992031591013
step: 20, loss: 0.0003480224695522338
step: 30, loss: 0.00013039256737101823
step: 40, loss: 0.00030376430368050933
step: 50, loss: 0.00031389627838507295
step: 60, loss: 0.0001547433785162866
step: 70, loss: 0.0001658012333791703
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```

```

step: 80, loss: 0.0012963830959051847
step: 90, loss: 0.00012188201071694493
step: 100, loss: 0.0003546375664882362
step: 110, loss: 0.0005544906598515809
step: 120, loss: 0.031229553744196892
step: 130, loss: 0.0005671890103258193
step: 140, loss: 0.00013419626338873059
step: 150, loss: 9.797692473512143e-05
step: 160, loss: 7.174241909524426e-05
step: 170, loss: 0.00022435169375967234
/usr/local/lib/python3.7/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=0.838709677419355, f1=0.8393285371702638,
best_f1=0.8530805687203792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004138823598623276
step: 10, loss: 0.0013879688922315836
step: 20, loss: 0.02529595047235489
step: 30, loss: 0.00019699726544786245
step: 40, loss: 0.00014352696598507464
step: 50, loss: 0.0003766919544432312
step: 60, loss: 0.006463534198701382
step: 70, loss: 0.00010393476986791939
step: 80, loss: 0.008344253525137901
step: 90, loss: 0.00021050659415777773
step: 100, loss: 0.00017896955250762403
step: 110, loss: 0.00016515971219632775
step: 120, loss: 0.0006534451968036592
step: 130, loss: 0.0017199073918163776
step: 140, loss: 0.00019172369502484798
step: 150, loss: 0.0003537582524586469
step: 160, loss: 0.00015123101184144616
step: 170, loss: 0.00010122109233634546
/usr/local/lib/python3.7/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=0.8379052369077307, f1=0.8481927710843373,
best_f1=0.8530805687203792
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, 238.89it/s]
load_f1 = 0.44444444444444445
real_f1 = 0.4430844553243574
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.39it/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: 0.6227690577507019
step: 10, loss: 0.6103795170783997
step: 20, loss: 0.314532071352005
step: 30, loss: 0.0945499986410141
step: 40, loss: 0.22383204102516174
step: 50, loss: 0.039627816528081894
step: 60, loss: 0.029441270977258682
step: 70, loss: 0.009375985711812973
step: 80, loss: 0.15081529319286346
step: 90, loss: 0.09985903650522232
step: 100, loss: 0.0033969818614423275
step: 110, loss: 0.16700947284698486
step: 120, loss: 0.020267223939299583
step: 130, loss: 0.051624927669763565
step: 140, loss: 0.0023172807414084673
step: 150, loss: 0.03179306164383888
step: 160, loss: 0.006034239660948515
step: 170, loss: 0.1428782343864441
step: 180, loss: 0.007277953904122114
step: 190, loss: 0.07736103981733322
step: 200, loss: 0.10574521124362946
step: 210, loss: 0.010329087264835835
step: 220, loss: 0.011055666953325272
step: 230, loss: 0.008407861925661564

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

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.976271186440678, f1=0.9749430523917996,
best_f1=0.9749430523917996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005171281285583973
step: 10, loss: 0.010063417255878448

```

step: 20, loss: 0.15455833077430725
step: 30, loss: 0.24215784668922424
step: 40, loss: 0.04298676177859306
step: 50, loss: 0.00807856023311615
step: 60, loss: 0.012433535419404507
step: 70, loss: 0.002893214114010334
step: 80, loss: 0.001823120517656207
step: 90, loss: 0.0035654595121741295
step: 100, loss: 0.04874570295214653
step: 110, loss: 0.14748281240463257
step: 120, loss: 0.12273797392845154
step: 130, loss: 0.03337007015943527
step: 140, loss: 0.002208269899711013
step: 150, loss: 0.006888079922646284
step: 160, loss: 0.0529693178832531
step: 170, loss: 0.0008759609772823751
step: 180, loss: 0.017231013625860214
step: 190, loss: 0.011526244692504406
step: 200, loss: 0.0016568463761359453
step: 210, loss: 0.0011818745406344533
step: 220, loss: 0.07820937782526016
step: 230, loss: 0.013589591719210148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, f1=0.9807474518686297,
best_f1=0.9807474518686297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004717702511698008
step: 10, loss: 0.0036757623311132193
step: 20, loss: 0.0036004262510687113
step: 30, loss: 0.006790100131183863
step: 40, loss: 0.10929574817419052
step: 50, loss: 0.003353598527610302
step: 60, loss: 0.003518315264955163
step: 70, loss: 0.0011626349296420813
step: 80, loss: 0.0009179091430269182
step: 90, loss: 0.003901897929608822
step: 100, loss: 0.0010625218274071813
step: 110, loss: 0.001710320240817964
step: 120, loss: 0.014037015847861767
step: 130, loss: 0.0007113935425877571

```

```

step: 140, loss: 0.0020256605930626392
step: 150, loss: 0.10606816411018372
step: 160, loss: 0.024330202490091324
step: 170, loss: 0.007429605349898338
step: 180, loss: 0.0032567186281085014
step: 190, loss: 0.0036791320890188217
step: 200, loss: 0.0008824457181617618
step: 210, loss: 0.001925349817611277
step: 220, loss: 0.0010338958818465471
step: 230, loss: 0.003912750165909529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9844097995545658, f1=0.9743016759776536,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000989715219475329
step: 10, loss: 0.0012935320846736431
step: 20, loss: 0.0006743580452166498
step: 30, loss: 0.0015670291613787413
step: 40, loss: 0.00981869362294674
step: 50, loss: 0.0008759775664657354
step: 60, loss: 0.0007128361612558365
step: 70, loss: 0.0007103480165824294
step: 80, loss: 0.005429306533187628
step: 90, loss: 0.004429454915225506
step: 100, loss: 0.00028478348394855857
step: 110, loss: 0.0018490799702703953
step: 120, loss: 0.03576665371656418
step: 130, loss: 0.0010858463356271386
step: 140, loss: 0.0015686546685174108
step: 150, loss: 0.05884955823421478
step: 160, loss: 0.014511079527437687
step: 170, loss: 0.26188403367996216
step: 180, loss: 0.0002211298851761967
step: 190, loss: 0.002842634217813611
step: 200, loss: 0.0005761734792031348
step: 210, loss: 0.0815930888056755
step: 220, loss: 0.0005634603439830244
step: 230, loss: 0.02308320440351963
/usr/local/lib/python3.7/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=0.9821029082774049, f1=0.9764309764309763,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003319552633911371
step: 10, loss: 0.00032570265466347337
step: 20, loss: 0.0005506580346263945
step: 30, loss: 0.0012302786344662309
step: 40, loss: 0.002474837703630328
step: 50, loss: 0.00021130738605279475
step: 60, loss: 0.13303090631961823
step: 70, loss: 0.0021389727480709553
step: 80, loss: 0.0021151932887732983
step: 90, loss: 0.0007866075611673295
step: 100, loss: 0.0005294416914694011
step: 110, loss: 0.004624588415026665
step: 120, loss: 0.00012144529318902642
step: 130, loss: 0.0011057807132601738
step: 140, loss: 0.0004337076097726822
step: 150, loss: 0.0011997136753052473
step: 160, loss: 0.005871493835002184
step: 170, loss: 0.011787962168455124
step: 180, loss: 0.03828691691160202
step: 190, loss: 0.005806622561067343
step: 200, loss: 0.0016607098514214158
step: 210, loss: 0.0018583836499601603
step: 220, loss: 0.00618829857558012
step: 230, loss: 0.0006980164907872677
/usr/local/lib/python3.7/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=0.978865406006674, f1=0.9731543624161074,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02693980559706688
step: 10, loss: 0.005943189840763807
step: 20, loss: 0.019893301650881767
step: 30, loss: 0.0012211723951622844
step: 40, loss: 0.0006703652907162905
step: 50, loss: 0.011891724541783333
step: 60, loss: 0.00029858900234103203
step: 70, loss: 0.0039373780600726604
step: 80, loss: 0.0006179905030876398
step: 90, loss: 0.0003552486887201667

```



```

step: 100, loss: 0.0807720422744751
step: 110, loss: 0.00029990149778313935
step: 120, loss: 0.00040830575744621456
step: 130, loss: 0.0003844580496661365
step: 140, loss: 0.0001371220132568851
step: 150, loss: 0.0418030209839344
step: 160, loss: 0.0006005369359627366
step: 170, loss: 0.00017778690380509943
step: 180, loss: 0.1472747027873993
step: 190, loss: 0.0007594221970066428
step: 200, loss: 0.0003681745147332549
step: 210, loss: 0.00029125026776455343
step: 220, loss: 0.12471701949834824
step: 230, loss: 0.027623435482382774
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9775280898876404,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04722157120704651
step: 10, loss: 0.0001260315184481442
step: 20, loss: 0.010098487138748169
step: 30, loss: 0.0002480585244484246
step: 40, loss: 0.0010826276848092675
step: 50, loss: 0.00029333840939216316
step: 60, loss: 0.000356811739038676
step: 70, loss: 0.024660607799887657
step: 80, loss: 0.0002556981344241649
step: 90, loss: 0.00013661755656357855
step: 100, loss: 7.414600986521691e-05
step: 110, loss: 0.0007670978084206581
step: 120, loss: 6.984394713072106e-05
step: 130, loss: 0.00037707953015342355
step: 140, loss: 0.04370652139186859
step: 150, loss: 0.001789834932424128
step: 160, loss: 0.06832627952098846
step: 170, loss: 0.00026828330010175705
step: 180, loss: 0.000377941585611552
step: 190, loss: 0.00011435552005423233
step: 200, loss: 0.024275217205286026
step: 210, loss: 6.055855560419336e-05
step: 220, loss: 0.002564653754234314
step: 230, loss: 0.0002827859134413302
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9786276715410572,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040699142846278846
step: 10, loss: 0.0008305227966047823
step: 20, loss: 0.0006657188641838729
step: 30, loss: 5.935468652751297e-05
step: 40, loss: 0.0011830901494249701
step: 50, loss: 0.00012685998808592558
step: 60, loss: 0.0002511480124667287
step: 70, loss: 0.017571259289979935
step: 80, loss: 8.337660983670503e-05
step: 90, loss: 0.0175996832549572
step: 100, loss: 0.0006413322407752275
step: 110, loss: 0.011396612972021103
step: 120, loss: 0.07532850652933121
step: 130, loss: 0.0006195480818860233
step: 140, loss: 0.00010764376202132553
step: 150, loss: 0.0001581064861966297
step: 160, loss: 0.00033319054637104273
step: 170, loss: 0.00029649099451489747
step: 180, loss: 0.005309441592544317
step: 190, loss: 0.0032050563022494316
step: 200, loss: 0.00010980257502524182
step: 210, loss: 0.0031511285342276096
step: 220, loss: 8.504542347509414e-05
step: 230, loss: 6.32123919785954e-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=0.9819819819819819, f1=0.976271186440678,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.6968078347854316e-05
step: 10, loss: 0.00035101122921332717
step: 20, loss: 0.005278774071484804
step: 30, loss: 0.0008280000183731318
step: 40, loss: 0.004548973403871059
step: 50, loss: 0.0001282433804590255
step: 60, loss: 6.719341763528064e-05
step: 70, loss: 3.6379969969857484e-05
step: 80, loss: 4.2630308598745614e-05
step: 90, loss: 0.0010111437877640128

```

```

step: 100, loss: 0.00032130625913850963
step: 110, loss: 0.00017920024401973933
step: 120, loss: 8.495692600263283e-05
step: 130, loss: 7.055661990307271e-05
step: 140, loss: 0.00012609281111508608
step: 150, loss: 0.08156076818704605
step: 160, loss: 0.00019152525055687875
step: 170, loss: 0.00047588901361450553
step: 180, loss: 0.0007049489067867398
step: 190, loss: 0.0006074814591556787
step: 200, loss: 0.0002449320745654404
step: 210, loss: 0.0002038813690887764
step: 220, loss: 7.11673274054192e-05
step: 230, loss: 0.0009973598644137383
/usr/local/lib/python3.7/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=0.9819004524886877, f1=0.9726651480637813,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006637350306846201
step: 10, loss: 3.493424446787685e-05
step: 20, loss: 5.201690146350302e-05
step: 30, loss: 5.716284067602828e-05
step: 40, loss: 5.1377595809753984e-05
step: 50, loss: 6.721771933371201e-05
step: 60, loss: 4.6130255213938653e-05
step: 70, loss: 0.000252228433964774
step: 80, loss: 4.163219637121074e-05
step: 90, loss: 0.00010059017222374678
step: 100, loss: 4.40041403635405e-05
step: 110, loss: 5.857060750713572e-05
step: 120, loss: 0.00010871874110307544
step: 130, loss: 7.764270412735641e-05
step: 140, loss: 0.034131571650505066
step: 150, loss: 0.013782541267573833
step: 160, loss: 0.0006519216112792492
step: 170, loss: 5.243842679192312e-05
step: 180, loss: 0.00013936676259618253
step: 190, loss: 0.013176549226045609
step: 200, loss: 3.9828039007261395e-05
step: 210, loss: 7.748341158730909e-05
step: 220, loss: 3.1492647394770756e-05
step: 230, loss: 0.0008377851918339729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853107344632768, f1=0.9738933030646991,
best_f1=0.9738933030646991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.737789302249439e-05
step: 10, loss: 5.6376546126557514e-05
step: 20, loss: 7.163952250266448e-05
step: 30, loss: 0.022869281470775604
step: 40, loss: 3.762405685847625e-05
step: 50, loss: 7.460596680175513e-05
step: 60, loss: 0.012050325982272625
step: 70, loss: 7.874625589465722e-05
step: 80, loss: 2.6929503292194568e-05
step: 90, loss: 2.8501448468887247e-05
step: 100, loss: 3.0881732527632266e-05
step: 110, loss: 0.0002460299292579293
step: 120, loss: 0.00013763933384325355
step: 130, loss: 2.3986616724869236e-05
step: 140, loss: 2.695554394449573e-05
step: 150, loss: 0.013554807752370834
step: 160, loss: 2.7592428523348644e-05
step: 170, loss: 0.006632156670093536
step: 180, loss: 0.0003480289306025952
step: 190, loss: 2.638184196257498e-05
step: 200, loss: 5.866649007657543e-05
step: 210, loss: 2.2690312107442878e-05
step: 220, loss: 2.7845835575135425e-05
step: 230, loss: 4.805112621397711e-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=0.984304932735426, f1=0.9751693002257337,
best_f1=0.9738933030646991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.060233939322643e-05
step: 10, loss: 0.0001338740548817441
step: 20, loss: 4.3720407120417804e-05
step: 30, loss: 9.493064862908795e-05
step: 40, loss: 6.31326183793135e-05
step: 50, loss: 0.00034150437568314373

```

```

step: 60, loss: 0.00016264028090517968
step: 70, loss: 5.0175505748484284e-05
step: 80, loss: 0.00019756289839278907
step: 90, loss: 2.3394362870021723e-05
step: 100, loss: 2.3290056560654193e-05
step: 110, loss: 2.6232944946968928e-05
step: 120, loss: 2.7160454919794574e-05
step: 130, loss: 2.871382821467705e-05
step: 140, loss: 9.756185318110511e-05
step: 150, loss: 6.750164175173268e-05
step: 160, loss: 2.8173717510071583e-05
step: 170, loss: 3.117585583822802e-05
step: 180, loss: 0.0035260911099612713
step: 190, loss: 9.247588604921475e-05
step: 200, loss: 1.8700682630878873e-05
step: 210, loss: 2.714935726544354e-05
step: 220, loss: 2.7461903300718404e-05
step: 230, loss: 0.02924022264778614
/usr/local/lib/python3.7/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=0.984304932735426, f1=0.9751693002257337,
best_f1=0.9738933030646991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.106535197934136e-05
step: 10, loss: 0.00017860699153970927
step: 20, loss: 2.9499948141165078e-05
step: 30, loss: 3.98137817683164e-05
step: 40, loss: 2.520838097552769e-05
step: 50, loss: 4.397770680952817e-05
step: 60, loss: 0.0003798733523581177
step: 70, loss: 2.286537164764013e-05
step: 80, loss: 2.480992225173395e-05
step: 90, loss: 0.0001697011321084574
step: 100, loss: 2.5591773010091856e-05
step: 110, loss: 0.0028432654216885567
step: 120, loss: 0.004354643169790506
step: 130, loss: 4.459445335669443e-05
step: 140, loss: 8.321108180098236e-05
step: 150, loss: 2.6173402147833258e-05
step: 160, loss: 5.8219520724378526e-05
step: 170, loss: 3.371497950865887e-05
step: 180, loss: 2.6389405320514925e-05
step: 190, loss: 2.1442327124532312e-05
step: 200, loss: 5.0590591854415834e-05
step: 210, loss: 1.7817817933973856e-05

```

```

step: 220, loss: 2.9525037462008186e-05
step: 230, loss: 7.230702613014728e-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=0.984304932735426, f1=0.9751693002257337,
best_f1=0.9738933030646991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035923137329518795
step: 10, loss: 0.0014866183046251535
step: 20, loss: 2.3137341486290097e-05
step: 30, loss: 3.2952462788671255e-05
step: 40, loss: 0.007601423654705286
step: 50, loss: 3.1023111660033464e-05
step: 60, loss: 2.4478316845488735e-05
step: 70, loss: 2.343161395401694e-05
step: 80, loss: 2.5975923563237302e-05
step: 90, loss: 2.5972094590542838e-05
step: 100, loss: 2.1479649149114266e-05
step: 110, loss: 0.00019441550830379128
step: 120, loss: 1.7542150089866482e-05
step: 130, loss: 2.3200695068226196e-05
step: 140, loss: 2.595355181256309e-05
step: 150, loss: 2.7327187126502395e-05
step: 160, loss: 0.006029793526977301
step: 170, loss: 3.584211663110182e-05
step: 180, loss: 4.371390969026834e-05
step: 190, loss: 4.973314207745716e-05
step: 200, loss: 2.2023361452738754e-05
step: 210, loss: 3.1257492082659155e-05
step: 220, loss: 1.7929534806171432e-05
step: 230, loss: 0.012994335032999516
/usr/local/lib/python3.7/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=0.984304932735426, f1=0.9751693002257337,
best_f1=0.9738933030646991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.832812631619163e-05
step: 10, loss: 1.646558303036727e-05
step: 20, loss: 2.5204692065017298e-05
step: 30, loss: 2.2306639948510565e-05
step: 40, loss: 6.29237329121679e-05
step: 50, loss: 0.02237321250140667

```

```

step: 60, loss: 6.677307101199403e-05
step: 70, loss: 5.905328725930303e-05
step: 80, loss: 7.083857053657994e-05
step: 90, loss: 7.729873323114589e-05
step: 100, loss: 2.1379060854087584e-05
step: 110, loss: 2.359905920457095e-05
step: 120, loss: 4.499523129197769e-05
step: 130, loss: 0.00030164848431013525
step: 140, loss: 2.0242910977685824e-05
step: 150, loss: 3.512798866722733e-05
step: 160, loss: 2.0350864360807464e-05
step: 170, loss: 1.846969098551199e-05
step: 180, loss: 3.854552414850332e-05
step: 190, loss: 2.7283054805593565e-05
step: 200, loss: 2.585642141639255e-05
step: 210, loss: 0.0007521308143623173
step: 220, loss: 7.308607746381313e-05
step: 230, loss: 2.073834002658259e-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=0.984304932735426, f1=0.9751693002257337,
best_f1=0.9738933030646991
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:14, 172.53it/s]
load_f1 = 0.9843400447427293
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 [00:22, 196.19it/s]
```

3.3.2 DBLP-GoogleScholar - Running the matcher

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


Downloading: 100% 433/433 [00:00<00:00, 377kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 249kB/s]
Downloading: 100% 440M/440M [00:06<00:00, 69.2MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

step: 0, loss: 0.6272138953208923
step: 10, loss: 0.5125640630722046
step: 20, loss: 0.44479724764823914
step: 30, loss: 0.04920768365263939
step: 40, loss: 0.19314727187156677
step: 50, loss: 0.08155731856822968
step: 60, loss: 0.05811131373047829
step: 70, loss: 0.055002134293317795
step: 80, loss: 0.09974237531423569

```
step: 90, loss: 0.08013369888067245
step: 100, loss: 0.006835054140537977
step: 110, loss: 0.05477404594421387
step: 120, loss: 0.07853231579065323
step: 130, loss: 0.10584671050310135
step: 140, loss: 0.08331869542598724
step: 150, loss: 0.03122691810131073
step: 160, loss: 0.08088690042495728
step: 170, loss: 0.1621633619070053
step: 180, loss: 0.03438882902264595
step: 190, loss: 0.008634121157228947
step: 200, loss: 0.1444578915834427
step: 210, loss: 0.070948027074337
step: 220, loss: 0.18481071293354034
step: 230, loss: 0.16814744472503662
step: 240, loss: 0.06304856389760971
step: 250, loss: 0.017387595027685165
step: 260, loss: 0.05840136483311653
step: 270, loss: 0.02325556054711342
step: 280, loss: 0.02687489427626133
step: 290, loss: 0.054232463240623474
step: 300, loss: 0.015711862593889236
step: 310, loss: 0.12035385519266129
step: 320, loss: 0.08202657848596573
step: 330, loss: 0.015482231043279171
step: 340, loss: 0.039504751563072205
step: 350, loss: 0.02144879661500454
step: 360, loss: 0.015127474442124367
step: 370, loss: 0.041497062891721725
step: 380, loss: 0.009472238831222057
step: 390, loss: 0.187191903591156
step: 400, loss: 0.18776537477970123
step: 410, loss: 0.05166742950677872
step: 420, loss: 0.012587178498506546
step: 430, loss: 0.1695055067539215
step: 440, loss: 0.03714223951101303
step: 450, loss: 0.0032675745896995068
step: 460, loss: 0.018979718908667564
step: 470, loss: 0.1587010771036148
step: 480, loss: 0.03868314251303673
step: 490, loss: 0.05222022533416748
step: 500, loss: 0.11036452651023865
step: 510, loss: 0.05150949954986572
step: 520, loss: 0.05868084356188774
step: 530, loss: 0.003500985214486718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9416126042632066, f1=0.9419354838709678,
best_f1=0.9419354838709678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2074434459209442
step: 10, loss: 0.052700623869895935
step: 20, loss: 0.004204926546663046
step: 30, loss: 0.017386838793754578
step: 40, loss: 0.1279171258211136
step: 50, loss: 0.06522808969020844
step: 60, loss: 0.004100855439901352
step: 70, loss: 0.04546058177947998
step: 80, loss: 0.06178569048643112
step: 90, loss: 0.011155079118907452
step: 100, loss: 0.005929466336965561
step: 110, loss: 0.004788652528077364
step: 120, loss: 0.13033753633499146
step: 130, loss: 0.042951736599206924
step: 140, loss: 0.03490333631634712
step: 150, loss: 0.018520498648285866
step: 160, loss: 0.008173634298145771
step: 170, loss: 0.007618471048772335
step: 180, loss: 0.04311525076627731
step: 190, loss: 0.10855922847986221
step: 200, loss: 0.007746195886284113
step: 210, loss: 0.03664647415280342
step: 220, loss: 0.08975192904472351
step: 230, loss: 0.004680528771132231
step: 240, loss: 0.004133810289204121
step: 250, loss: 0.016636330634355545
step: 260, loss: 0.0013927260879427195
step: 270, loss: 0.06000257655978203
step: 280, loss: 0.004277192987501621
step: 290, loss: 0.006784067954868078
step: 300, loss: 0.20255239307880402
step: 310, loss: 0.012488006614148617
step: 320, loss: 0.026398373767733574
step: 330, loss: 0.050489187240600586
step: 340, loss: 0.007782375440001488
step: 350, loss: 0.0006293224287219346
step: 360, loss: 0.022159622982144356
step: 370, loss: 0.063536137342453
step: 380, loss: 0.06595242768526077

```

```

step: 390, loss: 0.10128282755613327
step: 400, loss: 0.01952395588159561
step: 410, loss: 0.008094582706689835
step: 420, loss: 0.04452100768685341
step: 430, loss: 0.008332140743732452
step: 440, loss: 0.18526691198349
step: 450, loss: 0.020028170198202133
step: 460, loss: 0.07588507235050201
step: 470, loss: 0.12137224525213242
step: 480, loss: 0.22447508573532104
step: 490, loss: 0.02184591442346573
step: 500, loss: 0.10632895678281784
step: 510, loss: 0.008936961181461811
step: 520, loss: 0.07327268272638321
step: 530, loss: 0.010426423512399197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.945337620578778, f1=0.9445983379501385,
best_f1=0.9445983379501385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03469932824373245
step: 10, loss: 0.014881419949233532
step: 20, loss: 0.07972346991300583
step: 30, loss: 0.04308320954442024
step: 40, loss: 0.09767620265483856
step: 50, loss: 0.10891416668891907
step: 60, loss: 0.012922516092658043
step: 70, loss: 0.004104281309992075
step: 80, loss: 0.0009825011948123574
step: 90, loss: 0.005993726197630167
step: 100, loss: 0.08855164051055908
step: 110, loss: 0.0040559424087405205
step: 120, loss: 0.004188638646155596
step: 130, loss: 0.001213357667438686
step: 140, loss: 0.022434573620557785
step: 150, loss: 0.010843687690794468
step: 160, loss: 0.03320565074682236
step: 170, loss: 0.09347493946552277
step: 180, loss: 0.06744346767663956
step: 190, loss: 0.014717438258230686
step: 200, loss: 0.02734030969440937

```

```

step: 210, loss: 0.029683489352464676
step: 220, loss: 0.07892508804798126
step: 230, loss: 0.106305330991745
step: 240, loss: 0.001410230528563261
step: 250, loss: 0.029245661571621895
step: 260, loss: 0.005062737502157688
step: 270, loss: 0.0007138131186366081
step: 280, loss: 0.05263146013021469
step: 290, loss: 0.003326846519485116
step: 300, loss: 0.004726417828351259
step: 310, loss: 0.001146941795013845
step: 320, loss: 0.07965972274541855
step: 330, loss: 0.015463552437722683
step: 340, loss: 0.007479607127606869
step: 350, loss: 0.0054070730693638325
step: 360, loss: 0.02276005409657955
step: 370, loss: 0.012302057817578316
step: 380, loss: 0.019877037033438683
step: 390, loss: 0.021214311942458153
step: 400, loss: 0.012883353978395462
step: 410, loss: 0.00480011897161603
step: 420, loss: 0.09751330316066742
step: 430, loss: 0.005082877352833748
step: 440, loss: 0.11643733084201813
step: 450, loss: 0.07888588309288025
step: 460, loss: 0.07349886000156403
step: 470, loss: 0.013357178308069706
step: 480, loss: 0.0024049559142440557
step: 490, loss: 0.003184788627550006
step: 500, loss: 0.009093467146158218
step: 510, loss: 0.009860852733254433
step: 520, loss: 0.04774845018982887
step: 530, loss: 0.04771951958537102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9522006625650733, f1=0.9473189087488241,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006499642506241798
step: 10, loss: 0.0016214525094255805
step: 20, loss: 0.0022424582857638597

```

step: 30, loss: 0.009706858545541763
step: 40, loss: 0.002285362919792533
step: 50, loss: 0.003474823897704482
step: 60, loss: 0.0010906618554145098
step: 70, loss: 0.003270863089710474
step: 80, loss: 0.0043811919167637825
step: 90, loss: 0.04528234153985977
step: 100, loss: 0.03333690017461777
step: 110, loss: 0.04166316241025925
step: 120, loss: 0.00016911744023673236
step: 130, loss: 0.0005334029556252062
step: 140, loss: 0.004001638852059841
step: 150, loss: 0.03637410327792168
step: 160, loss: 0.0007637274102307856
step: 170, loss: 0.006811391096562147
step: 180, loss: 0.001365875476039946
step: 190, loss: 0.032010916620492935
step: 200, loss: 0.0015853194054216146
step: 210, loss: 0.05250735208392143
step: 220, loss: 0.0025759595446288586
step: 230, loss: 0.01324584148824215
step: 240, loss: 0.0016755807446315885
step: 250, loss: 0.01829802803695202
step: 260, loss: 0.0008136529941111803
step: 270, loss: 0.00194496416952461
step: 280, loss: 0.0035788246896117926
step: 290, loss: 0.009308384731411934
step: 300, loss: 0.0006052936660125852
step: 310, loss: 0.0023532977793365717
step: 320, loss: 0.009645102545619011
step: 330, loss: 0.003179140156134963
step: 340, loss: 0.0037523359060287476
step: 350, loss: 0.0309468861669302
step: 360, loss: 0.0034617455676198006
step: 370, loss: 0.000825301802251488
step: 380, loss: 0.003720508888363838
step: 390, loss: 0.0028700516559183598
step: 400, loss: 0.04267745465040207
step: 410, loss: 0.00830263365060091
step: 420, loss: 0.000210974074434489
step: 430, loss: 0.019337862730026245
step: 440, loss: 0.00032630388159304857
step: 450, loss: 0.003517633303999901
step: 460, loss: 0.0020201278384774923
step: 470, loss: 0.039286985993385315
step: 480, loss: 0.11440909653902054
step: 490, loss: 0.03176030144095421
step: 500, loss: 0.012409400194883347

```
step: 510, loss: 0.009759936481714249
step: 520, loss: 0.09275869280099869
step: 530, loss: 0.04594148322939873
/usr/local/lib/python3.7/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=0.947945205479452, f1=0.9406934306569343,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033232334535568953
step: 10, loss: 0.04394307732582092
step: 20, loss: 0.0004153251647949219
step: 30, loss: 0.10255979001522064
step: 40, loss: 0.0044686575420200825
step: 50, loss: 0.0007474993471987545
step: 60, loss: 0.2114517092704773
step: 70, loss: 0.0007913348963484168
step: 80, loss: 0.0012454527895897627
step: 90, loss: 0.001282778917811811
step: 100, loss: 0.0021158272866159678
step: 110, loss: 0.000295367615763098
step: 120, loss: 0.005309326108545065
step: 130, loss: 0.00020467456488404423
step: 140, loss: 0.0005149937351234257
step: 150, loss: 0.0006318918894976377
step: 160, loss: 0.0007282717269845307
step: 170, loss: 0.014415265060961246
step: 180, loss: 0.0018538719741627574
step: 190, loss: 0.002472598571330309
step: 200, loss: 0.0005991898360662162
step: 210, loss: 0.00454341433942318
step: 220, loss: 0.0006901614251546562
step: 230, loss: 0.0017528499010950327
step: 240, loss: 0.0006043786415830255
step: 250, loss: 0.00011142106814077124
step: 260, loss: 0.03171566501259804
step: 270, loss: 0.00014261584146879613
step: 280, loss: 0.00805355329066515
step: 290, loss: 0.14063899219036102
step: 300, loss: 0.014970945194363594
step: 310, loss: 0.0008779942290857434
step: 320, loss: 0.23256565630435944
step: 330, loss: 0.0345575213432312
step: 340, loss: 0.004045871552079916
step: 350, loss: 0.0025339543353766203
step: 360, loss: 0.02408246509730816
```

```

step: 370, loss: 0.002002442255616188
step: 380, loss: 0.0006579105393029749
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 390, loss: 0.00016488648543599993
step: 400, loss: 0.0028083904180675745
step: 410, loss: 0.00038113584741950035
step: 420, loss: 0.001269967877306044
step: 430, loss: 0.003032315755262971
step: 440, loss: 0.07577859610319138
step: 450, loss: 0.0049153706058859825
step: 460, loss: 0.003032551845535636
step: 470, loss: 0.030496375635266304
step: 480, loss: 0.00067197869066149
step: 490, loss: 0.0001875258603831753
step: 500, loss: 0.0015224542003124952
step: 510, loss: 0.02033783122897148
step: 520, loss: 0.01948395185172558
step: 530, loss: 0.0011331737041473389
/usr/local/lib/python3.7/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=0.9389671361502347, f1=0.9402427637721755,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000799915287643671
step: 10, loss: 0.00011960265692323446
step: 20, loss: 0.12990060448646545
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.0015161627670750022
step: 40, loss: 0.001363521907478571
step: 50, loss: 0.015679603442549706
step: 60, loss: 0.0008120599086396396
step: 70, loss: 0.0002635017444845289
step: 80, loss: 0.0030423523858189583
step: 90, loss: 9.838445112109184e-05
step: 100, loss: 0.0010103413369506598
step: 110, loss: 0.0003437153354752809
step: 120, loss: 0.041953738778829575
step: 130, loss: 0.001135500380769372
step: 140, loss: 0.0036346963606774807
step: 150, loss: 0.001229626708664
step: 160, loss: 0.0012734087649732828
step: 170, loss: 0.002816534135490656
step: 180, loss: 0.002351033501327038
step: 190, loss: 0.002106510568410158
step: 200, loss: 0.0016471069538965821

```



```

step: 210, loss: 0.010683617554605007
step: 220, loss: 0.001272866502404213
step: 230, loss: 0.0009404586162418127
step: 240, loss: 0.064890056848526
step: 250, loss: 0.0059693544171750546
step: 260, loss: 0.001191300107166171
step: 270, loss: 0.1018049567937851
step: 280, loss: 0.0012669612187892199
step: 290, loss: 0.0001816281146602705
step: 300, loss: 0.0017989813350141048
step: 310, loss: 0.0011062410194426775
step: 320, loss: 0.0005073543288744986
step: 330, loss: 0.00038257057894952595
step: 340, loss: 0.055787134915590286
step: 350, loss: 0.003338563023135066
step: 360, loss: 0.056964144110679626
step: 370, loss: 0.002619381295517087
step: 380, loss: 0.0006185296224430203
step: 390, loss: 0.024528656154870987
step: 400, loss: 0.0005204794579185545
step: 410, loss: 0.0008777563343755901
step: 420, loss: 0.02771705575287342
step: 430, loss: 0.003402675734832883
step: 440, loss: 0.000250757671892643
step: 450, loss: 0.00590133061632514
step: 460, loss: 0.0002485712175257504
step: 470, loss: 0.0015601952327415347
step: 480, loss: 0.004621237516403198
step: 490, loss: 0.00756089948117733
step: 500, loss: 0.0003235502808820456
step: 510, loss: 0.0044325487688183784
step: 520, loss: 0.0008559140260331333
step: 530, loss: 0.00675461208447814
/usr/local/lib/python3.7/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=0.9456572224802601, f1=0.9442379182156133,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001872822904260829
step: 10, loss: 0.00017416298214811832
step: 20, loss: 0.0004993964685127139
step: 30, loss: 0.0002411140303593129
step: 40, loss: 0.001900913892313838
step: 50, loss: 0.0005274125724099576
step: 60, loss: 0.0005402398528531194

```

step: 70, loss: 0.0001175268116639927
step: 80, loss: 0.0015612075803801417
step: 90, loss: 0.0004943122039549053
step: 100, loss: 5.344325109035708e-05
step: 110, loss: 0.00014634651597589254
step: 120, loss: 0.03919893875718117
step: 130, loss: 0.022012367844581604
step: 140, loss: 0.0003388917539268732
step: 150, loss: 0.003683351678773761
step: 160, loss: 0.0004319559666328132
step: 170, loss: 0.0003700072120409459
step: 180, loss: 0.00182827259413898
step: 190, loss: 0.002176789566874504
step: 200, loss: 0.0004968834109604359
step: 210, loss: 0.0003459598228801042
step: 220, loss: 0.00015996444562915713
step: 230, loss: 0.022370729595422745
step: 240, loss: 0.0012237434275448322
step: 250, loss: 0.0016607227735221386
step: 260, loss: 0.05333363637328148
step: 270, loss: 0.0002949370245914906
step: 280, loss: 0.016295133158564568
step: 290, loss: 0.00012787667219527066
step: 300, loss: 0.0009804972214624286
step: 310, loss: 6.940629828022793e-05
step: 320, loss: 9.360990952700377e-05
step: 330, loss: 0.0004005585506092757
step: 340, loss: 0.03154459595680237
step: 350, loss: 0.0009990752441808581
step: 360, loss: 0.0011597691336646676
step: 370, loss: 0.00021453008230309933
step: 380, loss: 0.0035648602060973644
step: 390, loss: 0.002541374647989869
step: 400, loss: 0.006070346105843782
step: 410, loss: 0.009165143594145775
step: 420, loss: 0.00034738852991722524
step: 430, loss: 0.00014281104085966945
step: 440, loss: 0.022254513576626778
step: 450, loss: 0.002161579206585884
step: 460, loss: 0.0008737319149076939
step: 470, loss: 0.001187796937301755
step: 480, loss: 0.041765209287405014
step: 490, loss: 0.00033481884747743607
step: 500, loss: 0.10429645329713821
step: 510, loss: 0.001672191545367241
step: 520, loss: 0.012298001907765865
step: 530, loss: 0.007167073432356119
/usr/local/lib/python3.7/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=0.9474169741697417, f1=0.9476609541454377,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007731271907687187
step: 10, loss: 0.001512480666860938
step: 20, loss: 0.000605534587521106
step: 30, loss: 0.02782372757792473
step: 40, loss: 0.00015179241017904133
step: 50, loss: 0.0018808523891493678
step: 60, loss: 0.0022910074330866337
step: 70, loss: 0.00015014596283435822
step: 80, loss: 0.0023373900912702084
step: 90, loss: 0.0012750799069181085
step: 100, loss: 0.0002160274307243526
step: 110, loss: 0.000766979530453682
step: 120, loss: 0.00015078761498443782
step: 130, loss: 0.0007729140343144536
step: 140, loss: 0.001541093341074884
step: 150, loss: 0.00023837154731154442
step: 160, loss: 0.012051618658006191
step: 170, loss: 0.015624983236193657
step: 180, loss: 4.8384325054939836e-05
step: 190, loss: 0.002092264359816909
step: 200, loss: 0.0008668207447044551
step: 210, loss: 0.005396413616836071
step: 220, loss: 0.003065513912588358
step: 230, loss: 0.00014822484808973968
step: 240, loss: 0.0035656243562698364
step: 250, loss: 7.106912380550057e-05
step: 260, loss: 8.06034222478047e-05
step: 270, loss: 0.0017033416079357266
step: 280, loss: 0.04866096377372742
step: 290, loss: 9.131299884757027e-05
step: 300, loss: 0.0008599628345109522
step: 310, loss: 0.01704380288720131
step: 320, loss: 0.005191300995647907
step: 330, loss: 0.0003390852943994105
step: 340, loss: 0.0002718045434448868
step: 350, loss: 9.243383829016238e-05
step: 360, loss: 0.004698093980550766
step: 370, loss: 0.0030116995330899954
step: 380, loss: 0.002622717758640647
step: 390, loss: 0.1169763058423996
step: 400, loss: 0.000513195525854826

```

```

step: 410, loss: 0.0007457372266799212
step: 420, loss: 0.002109767636284232
step: 430, loss: 0.038666997104883194
step: 440, loss: 0.00030021893326193094
step: 450, loss: 0.0006521873874589801
step: 460, loss: 0.00011553504009498283
step: 470, loss: 0.0002180763694923371
step: 480, loss: 5.346888428903185e-05
step: 490, loss: 0.006235154811292887
step: 500, loss: 0.00032639579148963094
step: 510, loss: 4.464192170416936e-05
step: 520, loss: 0.0022104838863015175
step: 530, loss: 0.0015087923966348171
/usr/local/lib/python3.7/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=0.9465861588481189, f1=0.943081906524757,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011353109730407596
step: 10, loss: 7.68107347539626e-05
step: 20, loss: 0.00468651857227087
step: 30, loss: 0.0005775913596153259
step: 40, loss: 7.20930693205446e-05
step: 50, loss: 8.255754801211879e-05
step: 60, loss: 5.1489900215528905e-05
step: 70, loss: 0.0008036339422687888
step: 80, loss: 0.00022874228307045996
step: 90, loss: 0.00041689147474244237
step: 100, loss: 0.00010494387242943048
step: 110, loss: 4.963374885846861e-05
step: 120, loss: 9.7026233561337e-05
step: 130, loss: 3.88752632716205e-05
step: 140, loss: 0.24817460775375366
step: 150, loss: 0.0002589416690170765
step: 160, loss: 0.0020895516499876976
step: 170, loss: 0.015446657314896584
step: 180, loss: 4.18202398577705e-05
step: 190, loss: 9.260045771952718e-05
step: 200, loss: 0.00015477300621569157
step: 210, loss: 8.731130947126076e-05
step: 220, loss: 0.015968192368745804
step: 230, loss: 0.0001708891213638708
step: 240, loss: 0.0006153101567178965
step: 250, loss: 0.001615517889149487
step: 260, loss: 0.01328076422214508

```

```

step: 270, loss: 3.706859206431545e-05
step: 280, loss: 6.368956383084878e-05
step: 290, loss: 0.038519442081451416
step: 300, loss: 0.00019414385315030813
step: 310, loss: 0.0035206840839236975
step: 320, loss: 0.0009799845283851027
step: 330, loss: 0.007809582632035017
step: 340, loss: 0.00011262496991548687
step: 350, loss: 0.009593253955245018
step: 360, loss: 2.6810230338014662e-05
step: 370, loss: 0.003297585528343916
step: 380, loss: 5.23595190315973e-05
step: 390, loss: 0.00010242244024993852
step: 400, loss: 5.7139255659421906e-05
step: 410, loss: 0.0004976129857823253
step: 420, loss: 3.3053383958758786e-05
step: 430, loss: 8.400656952289864e-05
step: 440, loss: 0.0021129215601831675
step: 450, loss: 8.163565507857129e-05
step: 460, loss: 5.680153481080197e-05
step: 470, loss: 2.2734977392246947e-05
step: 480, loss: 3.945273419958539e-05
step: 490, loss: 0.0010141517268493772
step: 500, loss: 0.0035240105353295803
step: 510, loss: 0.0024620762560516596
step: 520, loss: 0.0012594249565154314
step: 530, loss: 0.01947944052517414
/usr/local/lib/python3.7/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=0.9444184960298926, f1=0.93796992481203,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011002314276993275
step: 10, loss: 3.261697929701768e-05
step: 20, loss: 0.0013163211988285184
step: 30, loss: 3.573457070160657e-05
step: 40, loss: 0.00014463361003436148
step: 50, loss: 0.0009185195667669177
step: 60, loss: 0.00023168559710029513
step: 70, loss: 3.096694854320958e-05
step: 80, loss: 2.9864579119021073e-05
step: 90, loss: 7.026219827821478e-05
step: 100, loss: 4.310342046665028e-05
step: 110, loss: 2.079040314129088e-05
step: 120, loss: 3.717678919201717e-05

```

```

step: 130, loss: 5.2170351409586146e-05
step: 140, loss: 0.008502484299242496
step: 150, loss: 3.2363925129175186e-05
step: 160, loss: 3.700701563502662e-05
step: 170, loss: 7.664706208743155e-05
step: 180, loss: 0.00021426132298074663
step: 190, loss: 3.336236841278151e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.006650072056800127
step: 210, loss: 0.001820042496547103
step: 220, loss: 0.0005106726894155145
step: 230, loss: 0.0005888441810384393
step: 240, loss: 4.438676478457637e-05
step: 250, loss: 0.0003934673441108316
step: 260, loss: 0.0015853760996833444
step: 270, loss: 0.0008859446388669312
step: 280, loss: 0.00027858532848767936
step: 290, loss: 4.978281504008919e-05
step: 300, loss: 2.3327289454755373e-05
step: 310, loss: 0.00029340863693505526
step: 320, loss: 0.0015012981602922082
step: 330, loss: 6.0251397371757776e-05
step: 340, loss: 0.0013553719036281109
step: 350, loss: 5.20069406775292e-05
step: 360, loss: 6.858864799141884e-05
step: 370, loss: 0.0012713688192889094
step: 380, loss: 0.0006545344949699938
step: 390, loss: 0.00032162893330678344
step: 400, loss: 0.003341286676004529
step: 410, loss: 0.0025584797840565443
step: 420, loss: 3.666851989692077e-05
step: 430, loss: 2.063021202047821e-05
step: 440, loss: 0.0003013384121004492
step: 450, loss: 2.832613063219469e-05
step: 460, loss: 2.2328966224449687e-05
step: 470, loss: 0.00023644721659366041
step: 480, loss: 3.1198360375128686e-05
step: 490, loss: 7.842532795621082e-05
step: 500, loss: 0.0034002435859292746
step: 510, loss: 3.5408418625593185e-05
step: 520, loss: 5.693913044524379e-05
step: 530, loss: 0.0015138817252591252
/usr/local/lib/python3.7/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=0.9492619926199262, f1=0.945387792565397,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0004757156129926443
step: 10, loss: 5.851920650457032e-05
step: 20, loss: 3.599163028411567e-05
step: 30, loss: 7.074833411024883e-05
step: 40, loss: 3.777436359087005e-05
step: 50, loss: 0.00014333943545352668
step: 60, loss: 0.07043591886758804
step: 70, loss: 6.909652438480407e-05
step: 80, loss: 0.00011972121865255758
step: 90, loss: 6.488407962024212e-05
step: 100, loss: 0.0008393483003601432
step: 110, loss: 0.00018384368740953505
step: 120, loss: 2.787520679703448e-05
step: 130, loss: 1.8812284906744026e-05
step: 140, loss: 0.0004980769590474665
step: 150, loss: 2.032843985944055e-05
step: 160, loss: 2.9387596441665664e-05
step: 170, loss: 3.40327387675643e-05
step: 180, loss: 3.339599061291665e-05
step: 190, loss: 9.258541831513867e-05
step: 200, loss: 0.0008219715673476458
step: 210, loss: 4.6551580453524366e-05
step: 220, loss: 0.0024153776466846466
step: 230, loss: 1.6081816283985972e-05
step: 240, loss: 0.0003713062033057213
step: 250, loss: 0.003595219226554036
step: 260, loss: 6.463242607424036e-05
step: 270, loss: 9.252899326384068e-05
step: 280, loss: 3.945583739550784e-05
step: 290, loss: 4.4696120312437415e-05
step: 300, loss: 2.7342694011167623e-05
step: 310, loss: 5.199079532758333e-05
step: 320, loss: 0.0006364543223753572
step: 330, loss: 0.0003801556886173785
step: 340, loss: 0.00015510720550082624
step: 350, loss: 9.555973520036787e-05
step: 360, loss: 4.4752025132765993e-05
step: 370, loss: 0.0009889150969684124
step: 380, loss: 2.2414666091208346e-05
step: 390, loss: 4.660476770368405e-05
step: 400, loss: 3.191966607118957e-05
step: 410, loss: 4.5069238694850355e-05
step: 420, loss: 7.625889702467248e-05
step: 430, loss: 6.746129656676203e-05
step: 440, loss: 0.00015683397941756994
step: 450, loss: 0.000741190102417022
```

```

step: 460, loss: 0.00242939917370677
step: 470, loss: 2.8631402528844774e-05
step: 480, loss: 0.0006382434512488544
step: 490, loss: 2.2503787477035075e-05
step: 500, loss: 4.826689837500453e-05
step: 510, loss: 4.736415212391876e-05
step: 520, loss: 3.0855026125209406e-05
step: 530, loss: 3.2843472581589594e-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=0.9506517690875234, f1=0.9484440315838365,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001138691877713427
step: 10, loss: 2.3338281607721e-05
step: 20, loss: 0.00014197167183738202
step: 30, loss: 0.0010483458172529936
step: 40, loss: 0.0004931040457449853
step: 50, loss: 0.02913973107933998
step: 60, loss: 0.0015152761479839683
step: 70, loss: 3.932559411623515e-05
step: 80, loss: 2.5319779524579644e-05
step: 90, loss: 4.1036208131117746e-05
step: 100, loss: 0.0004603570559993386
step: 110, loss: 3.790978007600643e-05
step: 120, loss: 3.440247746766545e-05
step: 130, loss: 0.0003457940765656531
step: 140, loss: 2.882061744458042e-05
step: 150, loss: 2.4536844648537226e-05
step: 160, loss: 1.885337769635953e-05
step: 170, loss: 2.9714530683122575e-05
step: 180, loss: 3.479100632830523e-05
step: 190, loss: 0.00021434749942272902
step: 200, loss: 0.00011911919864360243
step: 210, loss: 5.393888568505645e-05
step: 220, loss: 2.913842399721034e-05
step: 230, loss: 3.466755515546538e-05
step: 240, loss: 9.71213448792696e-05
step: 250, loss: 2.6042644094559364e-05
step: 260, loss: 0.00021668725821655244
step: 270, loss: 7.708312477916479e-05
step: 280, loss: 0.0009370243642479181
step: 290, loss: 0.000109152402728796
step: 300, loss: 0.0007747196359559894
step: 310, loss: 0.0018507607746869326

```



```

step: 320, loss: 5.9815611166413873e-05
step: 330, loss: 0.000225158961256966
step: 340, loss: 4.8096968384925276e-05
step: 350, loss: 0.00018563185585662723
step: 360, loss: 6.363770080497488e-05
step: 370, loss: 6.330099131446332e-05
step: 380, loss: 0.00026045911363326013
step: 390, loss: 0.00013761970330961049
step: 400, loss: 4.278264532331377e-05
step: 410, loss: 0.00014410674339160323
step: 420, loss: 0.003129089018329978
step: 430, loss: 0.025380482897162437
step: 440, loss: 0.00022292582434602082
step: 450, loss: 0.00044770858949050307
step: 460, loss: 2.9175404051784426e-05
step: 470, loss: 0.00012257102935109288
step: 480, loss: 0.0002668576780706644
step: 490, loss: 0.0016050497069954872
step: 500, loss: 4.302244269638322e-05
step: 510, loss: 0.004862278699874878
step: 520, loss: 4.1266339394496754e-05
step: 530, loss: 0.0011512566125020385
/usr/local/lib/python3.7/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=0.9472209248014946, f1=0.9420491423273065,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002539761131629348
step: 10, loss: 0.0001990277523873374
step: 20, loss: 0.0006335433572530746
step: 30, loss: 5.280123514239676e-05
step: 40, loss: 3.152565113850869e-05
step: 50, loss: 3.8688092899974436e-05
step: 60, loss: 0.000456265959655866
step: 70, loss: 2.396773561486043e-05
step: 80, loss: 5.526113818632439e-05
step: 90, loss: 0.03515224531292915
step: 100, loss: 0.0012393792858347297
step: 110, loss: 7.01817189110443e-05
step: 120, loss: 9.836532262852415e-05
step: 130, loss: 4.5531054638559e-05
step: 140, loss: 0.00010512829612707719
step: 150, loss: 8.258574962383136e-05
step: 160, loss: 3.524070052662864e-05
step: 170, loss: 8.09052144177258e-05

```

```

step: 180, loss: 5.039390816818923e-05
step: 190, loss: 0.00019851977413054556
step: 200, loss: 2.5372357413289137e-05
step: 210, loss: 2.952171598735731e-05
step: 220, loss: 2.1300716980476864e-05
step: 230, loss: 2.598968603706453e-05
step: 240, loss: 6.75206320011057e-05
step: 250, loss: 2.943950858025346e-05
step: 260, loss: 3.452667806413956e-05
step: 270, loss: 1.9330123905092478e-05
step: 280, loss: 4.276847903383896e-05
step: 290, loss: 4.9158345063915476e-05
step: 300, loss: 0.00013029511319473386
step: 310, loss: 0.0017151921056210995
step: 320, loss: 0.0028256624937057495
step: 330, loss: 0.001706751761958003
step: 340, loss: 0.0002679148456081748
step: 350, loss: 5.611423694062978e-05
step: 360, loss: 2.3785421944921836e-05
step: 370, loss: 3.987184027209878e-05
step: 380, loss: 0.00017586634203325957
step: 390, loss: 0.003942993935197592
step: 400, loss: 3.056077548535541e-05
step: 410, loss: 0.0008291739504784346
step: 420, loss: 2.5040570108103566e-05
step: 430, loss: 2.3714583221590146e-05
step: 440, loss: 6.179649790283293e-05
step: 450, loss: 8.839467773213983e-05
step: 460, loss: 0.0009353397181257606
step: 470, loss: 2.8209073207108304e-05
step: 480, loss: 0.00016442961350549012
step: 490, loss: 2.3614034944330342e-05
step: 500, loss: 0.00016004205099307
step: 510, loss: 4.1390288970433176e-05
step: 520, loss: 0.000412770634284243
step: 530, loss: 3.3162239560624585e-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=0.9498141263940519, f1=0.9459084604715674,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010089670540764928
step: 10, loss: 2.0651723389164545e-05
step: 20, loss: 4.07309053116478e-05
step: 30, loss: 0.0009462754242122173

```

step: 40, loss: 0.0002871341130230576
step: 50, loss: 0.01188327744603157
step: 60, loss: 2.6939560484606773e-05
step: 70, loss: 0.00017382750229444355
step: 80, loss: 6.858209235360846e-05
step: 90, loss: 4.0069262468023226e-05
step: 100, loss: 0.00027595239225775003
step: 110, loss: 0.00016384570335503668
step: 120, loss: 5.798431811854243e-05
step: 130, loss: 0.015385101549327374
step: 140, loss: 0.0016711711650714278
step: 150, loss: 2.27980508498149e-05
step: 160, loss: 0.0021561444737017155
step: 170, loss: 6.20586724835448e-05
step: 180, loss: 8.92354582902044e-05
step: 190, loss: 2.2731117496732622e-05
step: 200, loss: 8.291035192087293e-05
step: 210, loss: 3.062071846215986e-05
step: 220, loss: 3.68135588360019e-05
step: 230, loss: 0.0016806380590423942
step: 240, loss: 0.00017648683569859713
step: 250, loss: 3.986623414675705e-05
step: 260, loss: 0.0010397135047242045
step: 270, loss: 0.0007197035592980683
step: 280, loss: 2.3308391973841935e-05
step: 290, loss: 0.0013976009795442224
step: 300, loss: 0.00019634129421319813
step: 310, loss: 6.123050843598321e-05
step: 320, loss: 0.0016897546593099833
step: 330, loss: 8.683014311827719e-05
step: 340, loss: 0.0001311007799813524
step: 350, loss: 9.136854350799695e-05
step: 360, loss: 7.366808131337166e-05
step: 370, loss: 2.8292655770201236e-05
step: 380, loss: 2.6303536287741736e-05
step: 390, loss: 0.029498150572180748
step: 400, loss: 0.0016616801731288433
step: 410, loss: 2.3952692572493106e-05
step: 420, loss: 8.218201401177794e-05
step: 430, loss: 2.6728313969215378e-05
step: 440, loss: 6.782125274185091e-05
step: 450, loss: 0.005183254834264517
step: 460, loss: 0.0011372605804353952
step: 470, loss: 0.00019140368385706097
step: 480, loss: 2.0283756384742446e-05
step: 490, loss: 2.4030920030782e-05
step: 500, loss: 1.862606222857721e-05
step: 510, loss: 5.650894672726281e-05

```

step: 520, loss: 0.000413833447964862
step: 530, loss: 5.5380165576934814e-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=0.9489322191272052, f1=0.9453703703703704,
best_f1=0.9473189087488241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7651498385239393e-05
step: 10, loss: 3.267520514782518e-05
step: 20, loss: 4.4288241042522714e-05
step: 30, loss: 6.765830039512366e-05
step: 40, loss: 0.06947829574346542
step: 50, loss: 0.00019178497313987464
step: 60, loss: 1.436445745639503e-05
step: 70, loss: 2.902495180023834e-05
step: 80, loss: 3.10079449263867e-05
step: 90, loss: 2.622161991894245e-05
step: 100, loss: 8.282416820293292e-05
step: 110, loss: 0.0019489957485347986
step: 120, loss: 2.9905146220698953e-05
step: 130, loss: 0.10833068937063217
step: 140, loss: 3.9691665733698756e-05
step: 150, loss: 3.752984048333019e-05
step: 160, loss: 2.837029387592338e-05
step: 170, loss: 2.882403896364849e-05
step: 180, loss: 2.1508843929041177e-05
step: 190, loss: 2.5547495170030743e-05
step: 200, loss: 2.2947158868191764e-05
step: 210, loss: 0.0005029112217016518
step: 220, loss: 4.312819874030538e-05
step: 230, loss: 6.617964390898123e-05
step: 240, loss: 2.168815626646392e-05
step: 250, loss: 4.5878408855060115e-05
step: 260, loss: 2.52900026680436e-05
step: 270, loss: 2.6031326342490502e-05
step: 280, loss: 3.6498968256637454e-05
step: 290, loss: 1.8629742044140585e-05
step: 300, loss: 2.5539293346810155e-05
step: 310, loss: 5.226975918048993e-05
step: 320, loss: 5.253514609648846e-05
step: 330, loss: 5.751403296017088e-05
step: 340, loss: 2.346485598536674e-05
step: 350, loss: 1.5899089703452773e-05
step: 360, loss: 0.0011467082658782601
step: 370, loss: 1.42341423270409e-05

```

```

step: 380, loss: 0.00017197031411342323
step: 390, loss: 0.00010289695637766272
step: 400, loss: 2.0712184777949005e-05
step: 410, loss: 3.0149440135573968e-05
step: 420, loss: 4.342526153777726e-05
step: 430, loss: 3.533663766575046e-05
step: 440, loss: 0.0001825619110604748
step: 450, loss: 1.4882306459185202e-05
step: 460, loss: 2.5681103579699993e-05
step: 470, loss: 0.002180564682930708
step: 480, loss: 1.4319834008347243e-05
step: 490, loss: 1.826098741730675e-05
step: 500, loss: 2.293589750479441e-05
step: 510, loss: 3.373140498297289e-05
step: 520, loss: 1.9903844076907262e-05
step: 530, loss: 0.00039867713348940015
/usr/local/lib/python3.7/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=0.9488847583643123, f1=0.947075208913649,
best_f1=0.9473189087488241
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 [00:31, 182.98it/s]
load_f1 = 0.9521152952115295
real_f1 = 0.9517625231910947
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, 176.55it/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: 0.5536436438560486

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

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

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

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=0.5405405405405405, f1=0.4444444444444444,
best_f1=0.4444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4885057210922241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7142857142857143, 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: 0.3981545567512512
/usr/local/lib/python3.7/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=0.6250000000000001, f1=0.4864864864864865,
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: 0.35779622197151184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.6250000000000001,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2640339434146881
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.7142857142857143,
best_f1=0.6250000000000001

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24227085709571838
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.7857142857142857,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03904504328966141
/usr/local/lib/python3.7/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=0.8235294117647058, f1=0.8, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01579519733786583
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.8387096774193549,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016141717787832022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8750000000000001, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025805732235312462
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8125000000000001,
best_f1=0.8484848484848484

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013970609288662672
/usr/local/lib/python3.7/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=0.8750000000000001, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028325822204351425
/usr/local/lib/python3.7/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=0.8750000000000001, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017487421864643693
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1186344251036644
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001956785563379526
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.8484848484848484,
best_f1=0.8484848484848484
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, 137049.07it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7368421052631579
real_f1 = 0.717948717948718
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, 254.27it/s]

```

4.1.2 DBLP-ACM - Running the matcher

```

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

```

```

--n_epochs 15 \
--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5973193645477295
step: 10, loss: 0.6070447564125061
step: 20, loss: 0.26866111159324646
step: 30, loss: 0.1015532836318016
step: 40, loss: 0.20412634313106537
step: 50, loss: 0.0155104361474514
step: 60, loss: 0.16266527771949768
step: 70, loss: 0.06094525009393692
step: 80, loss: 0.028401440009474754
step: 90, loss: 0.16919536888599396
step: 100, loss: 0.009742497466504574
step: 110, loss: 0.2071951925754547
step: 120, loss: 0.021708769723773003
step: 130, loss: 0.007910741493105888
step: 140, loss: 0.026487646624445915
step: 150, loss: 0.006414094008505344
step: 160, loss: 0.004911113064736128
step: 170, loss: 0.02730623632669449
step: 180, loss: 0.0019584447145462036
step: 190, loss: 0.03325521945953369
step: 200, loss: 0.005830823909491301
step: 210, loss: 0.017118249088525772
step: 220, loss: 0.005753674544394016
step: 230, loss: 0.0038024443201720715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.9829351535836178,
best_f1=0.9829351535836178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007433232967741787
step: 10, loss: 0.0009469612268730998
step: 20, loss: 0.06762144714593887
step: 30, loss: 0.16570189595222473
step: 40, loss: 0.09214161336421967
step: 50, loss: 0.004163036122918129
step: 60, loss: 0.003825059160590172
step: 70, loss: 0.04185482859611511
step: 80, loss: 0.001976762665435672
step: 90, loss: 0.014015326276421547
step: 100, loss: 0.014286470599472523
step: 110, loss: 0.03944395110011101
step: 120, loss: 0.11644008755683899
step: 130, loss: 0.055929865688085556
step: 140, loss: 0.0016371896490454674
step: 150, loss: 0.0020832500886172056
step: 160, loss: 0.005642611533403397
step: 170, loss: 0.0009736213833093643
step: 180, loss: 0.0038384245708584785
step: 190, loss: 0.0021217670291662216
step: 200, loss: 0.0010373818222433329
step: 210, loss: 0.0006643523229286075
step: 220, loss: 0.051603883504867554
step: 230, loss: 0.001428291667252779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932432432432432, f1=0.9886877828054299,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006985901854932308
step: 10, loss: 0.0032925368286669254
step: 20, loss: 0.00041099669761024415
step: 30, loss: 0.0157155804336071
step: 40, loss: 0.11513163894414902
step: 50, loss: 0.0093535166233778
step: 60, loss: 0.0027724362444132566
step: 70, loss: 0.004472628235816956
step: 80, loss: 0.0027327705174684525
step: 90, loss: 0.007689111866056919

```

```

step: 100, loss: 0.0023083684500306845
step: 110, loss: 0.0003788692702073604
step: 120, loss: 0.0052959611639380455
step: 130, loss: 0.0012050438672304153
step: 140, loss: 0.0011684064520522952
step: 150, loss: 0.0077324179001152515
step: 160, loss: 0.004935829900205135
step: 170, loss: 0.007493453100323677
step: 180, loss: 0.003341112518683076
step: 190, loss: 0.0015498081920668483
step: 200, loss: 0.019280267879366875
step: 210, loss: 0.0002966580796055496
step: 220, loss: 0.00016433930431958288
step: 230, loss: 0.00014440368977375329
/usr/local/lib/python3.7/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=0.9921436588103255, f1=0.9821428571428571,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002501104318071157
step: 10, loss: 0.0001485128013882786
step: 20, loss: 0.0011518674436956644
step: 30, loss: 0.008280755952000618
step: 40, loss: 0.0007656746893189847
step: 50, loss: 0.00033594848355278373
step: 60, loss: 0.00021529443620238453
step: 70, loss: 0.0002712125133257359
step: 80, loss: 0.019933553412556648
step: 90, loss: 0.0005704692448489368
step: 100, loss: 0.00012989205424673855
step: 110, loss: 0.00017973105423152447
step: 120, loss: 0.036973029375076294
step: 130, loss: 0.00036328317946754396
step: 140, loss: 0.0001034121960401535
step: 150, loss: 0.04625188931822777
step: 160, loss: 0.00017952309281099588
step: 170, loss: 0.000554910278879106
step: 180, loss: 0.00012305496784392744
step: 190, loss: 0.002938157180324197
step: 200, loss: 0.0007376350113190711
step: 210, loss: 0.15325352549552917
step: 220, loss: 0.0007354571716859937
step: 230, loss: 0.002505512908101082
/usr/local/lib/python3.7/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=0.9886877828054299, f1=0.9830890642615557,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0041219089180231094
step: 10, loss: 0.001959142740815878
step: 20, loss: 0.003013626206666231
step: 30, loss: 0.00011852916213683784
step: 40, loss: 0.0001298470888286829
step: 50, loss: 0.0001643847208470106
step: 60, loss: 0.03245458006858826
step: 70, loss: 0.001744510605931282
step: 80, loss: 0.006064117420464754
step: 90, loss: 0.000391397945350036
step: 100, loss: 0.012947975657880306
step: 110, loss: 0.0005806654808111489
step: 120, loss: 0.0007979278452694416
step: 130, loss: 0.0038995135109871626
step: 140, loss: 0.0005859181401319802
step: 150, loss: 0.001886510755866766
step: 160, loss: 0.0009578344179317355
step: 170, loss: 0.013068869709968567
step: 180, loss: 0.0031739927362650633
step: 190, loss: 0.06396040320396423
step: 200, loss: 0.0009038621792569757
step: 210, loss: 0.0005068301688879728
step: 220, loss: 0.0004518183704931289
step: 230, loss: 0.0002938981051556766
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.9865470852017937,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004474678076803684
step: 10, loss: 0.0001698186679277569
step: 20, loss: 0.0023211254738271236
step: 30, loss: 0.00019020050240214914
step: 40, loss: 0.00033963884925469756
step: 50, loss: 0.0010132029419764876
step: 60, loss: 0.0003412803926039487
step: 70, loss: 0.0006202527438290417
step: 80, loss: 0.00027549167862161994
step: 90, loss: 0.012837782502174377

```



```

step: 100, loss: 0.03869669884443283
step: 110, loss: 0.008534637279808521
step: 120, loss: 0.0015212476719170809
step: 130, loss: 0.000341783365001902
step: 140, loss: 8.746290404815227e-05
step: 150, loss: 0.004150996450334787
step: 160, loss: 0.0012077372521162033
step: 170, loss: 0.00023207772755995393
step: 180, loss: 0.04521297663450241
step: 190, loss: 0.05752618983387947
step: 200, loss: 6.805211160099134e-05
step: 210, loss: 0.00019720832642633468
step: 220, loss: 7.286281470442191e-05
step: 230, loss: 0.056003499776124954
/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9876265466816648,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003858037758618593
step: 10, loss: 5.6152897741412744e-05
step: 20, loss: 0.009079552255570889
step: 30, loss: 0.0160201508551836
step: 40, loss: 0.00226930296048522
step: 50, loss: 0.00031447908258996904
step: 60, loss: 0.07470674812793732
step: 70, loss: 0.024254491552710533
step: 80, loss: 0.0002129426138708368
step: 90, loss: 0.0002084481529891491
step: 100, loss: 6.73458052915521e-05
step: 110, loss: 0.000182350559043698
step: 120, loss: 0.00019429206440690905
step: 130, loss: 3.6335575714474544e-05
step: 140, loss: 7.049093983368948e-05
step: 150, loss: 0.006076473277062178
step: 160, loss: 0.06200532615184784
step: 170, loss: 0.1277228146791458
step: 180, loss: 0.001030589104630053
step: 190, loss: 0.0001045488243107684
step: 200, loss: 0.05236422270536423
step: 210, loss: 3.395531894057058e-05
step: 220, loss: 0.010530216619372368
step: 230, loss: 0.00014952568744774908
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.984304932735426,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.731528340722434e-05
step: 10, loss: 0.008245494216680527
step: 20, loss: 8.951629570219666e-05
step: 30, loss: 4.781199459102936e-05
step: 40, loss: 0.0003064563497900963
step: 50, loss: 0.00012891656660940498
step: 60, loss: 3.866722545353696e-05
step: 70, loss: 6.70484805596061e-05
step: 80, loss: 2.8814514735131525e-05
step: 90, loss: 3.954440398956649e-05
step: 100, loss: 0.0001567777944728732
step: 110, loss: 0.06798052042722702
step: 120, loss: 0.0003748747694771737
step: 130, loss: 0.003676858963444829
step: 140, loss: 8.424321276834235e-05
step: 150, loss: 0.0020652473904192448
step: 160, loss: 0.0004835914296563715
step: 170, loss: 8.926819282351062e-05
step: 180, loss: 0.0003336384252179414
step: 190, loss: 0.00013354781549423933
step: 200, loss: 5.490767944138497e-05
step: 210, loss: 4.76818167953752e-05
step: 220, loss: 0.000135974187287502
step: 230, loss: 4.700954013969749e-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=0.9910514541387023, f1=0.9800884955752212,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.386286208406091e-05
step: 10, loss: 9.101913019549102e-05
step: 20, loss: 0.00018991442630067468
step: 30, loss: 0.0008661977481096983
step: 40, loss: 0.13332180678844452
step: 50, loss: 0.0002032900374615565
step: 60, loss: 0.00020789401605725288
step: 70, loss: 0.0001465399982407689
step: 80, loss: 0.00021022529108449817
step: 90, loss: 9.281462553190067e-05

```

```

step: 100, loss: 0.00017952789494302124
step: 110, loss: 6.788993050577119e-05
step: 120, loss: 0.00020703226618934423
step: 130, loss: 0.0007450571283698082
step: 140, loss: 3.569439286366105e-05
step: 150, loss: 0.0010725437896326184
step: 160, loss: 9.044590115081519e-05
step: 170, loss: 0.0001691766083240509
step: 180, loss: 0.001203565625473857
step: 190, loss: 0.00018925993936136365
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 200, loss: 0.0006934747216291726
step: 210, loss: 0.00020688031509052962
step: 220, loss: 3.8514066545758396e-05
step: 230, loss: 0.00013165500422473997
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9841269841269841,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.719781878113281e-05
step: 10, loss: 5.0169735914096236e-05
step: 20, loss: 0.001008627936244011
step: 30, loss: 4.111300950171426e-05
step: 40, loss: 0.00010304214083589613
step: 50, loss: 3.892872700816952e-05
step: 60, loss: 0.0006202380754984915
step: 70, loss: 0.00012226052058394998
step: 80, loss: 3.2297379220835865e-05
step: 90, loss: 0.0001812698901630938
step: 100, loss: 4.486594480113126e-05
step: 110, loss: 0.00010633302736096084
step: 120, loss: 0.0014332978753373027
step: 130, loss: 4.441010605660267e-05
step: 140, loss: 0.03212004899978638
step: 150, loss: 0.01687171496450901
step: 160, loss: 2.4895414753700607e-05
step: 170, loss: 0.00031991556170396507
step: 180, loss: 0.00011474924394860864
step: 190, loss: 0.017981503158807755
step: 200, loss: 3.254405964980833e-05
step: 210, loss: 0.0003518509038258344
step: 220, loss: 3.089617894147523e-05
step: 230, loss: 0.0002320724306628108
/usr/local/lib/python3.7/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=0.9921436588103255, f1=0.9865771812080537,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.837195981759578e-05
step: 10, loss: 8.192507084459066e-05
step: 20, loss: 3.169682531733997e-05
step: 30, loss: 0.012090750969946384
step: 40, loss: 0.0008457814692519605
step: 50, loss: 2.78754741884768e-05
step: 60, loss: 0.0006593568250536919
step: 70, loss: 0.0032855861354619265
step: 80, loss: 3.3492015063529834e-05
step: 90, loss: 6.895683327456936e-05
step: 100, loss: 2.4057311748038046e-05
step: 110, loss: 0.02171744592487812
step: 120, loss: 2.760689312708564e-05
step: 130, loss: 2.070467235171236e-05
step: 140, loss: 7.460286724381149e-05
step: 150, loss: 0.04756725952029228
step: 160, loss: 2.1297075363690965e-05
step: 170, loss: 0.020066944882273674
step: 180, loss: 8.816440822556615e-05
step: 190, loss: 2.6683195756049827e-05
step: 200, loss: 4.5548385969595984e-05
step: 210, loss: 1.862246972450521e-05
step: 220, loss: 2.3673614123254083e-05
step: 230, loss: 2.393061004113406e-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=0.9921436588103255, f1=0.9831649831649831,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.811391641444061e-05
step: 10, loss: 2.0861083612544462e-05
step: 20, loss: 2.87091334030265e-05
step: 30, loss: 0.0005451214965432882
step: 40, loss: 5.4531203204533085e-05
step: 50, loss: 5.944010990788229e-05
step: 60, loss: 0.00015789332974236459
step: 70, loss: 4.486609759624116e-05
step: 80, loss: 0.0002996120892930776

```

```

step: 90, loss: 1.6625803255010396e-05
step: 100, loss: 2.260433575429488e-05
step: 110, loss: 2.476486588420812e-05
step: 120, loss: 3.536923759384081e-05
step: 130, loss: 2.3409327695844695e-05
step: 140, loss: 0.01348070614039898
step: 150, loss: 5.9820886235684156e-05
step: 160, loss: 2.5297365937149152e-05
step: 170, loss: 0.000409166153986007
step: 180, loss: 0.03652924671769142
step: 190, loss: 0.0004974260227754712
step: 200, loss: 1.638363210076932e-05
step: 210, loss: 2.5431882022530772e-05
step: 220, loss: 2.1021347492933273e-05
step: 230, loss: 0.03884700685739517
/usr/local/lib/python3.7/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=0.990990990990991, f1=0.9864864864864865,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037615110632032156
step: 10, loss: 0.00020906671124976128
step: 20, loss: 0.000809025892522186
step: 30, loss: 4.4479591451818123e-05
step: 40, loss: 2.6333191271987744e-05
step: 50, loss: 0.00011044828715967014
step: 60, loss: 0.00018743185501080006
step: 70, loss: 1.4118715625954792e-05
step: 80, loss: 2.6324983991798945e-05
step: 90, loss: 0.00026321341283619404
step: 100, loss: 1.8134414858650416e-05
step: 110, loss: 0.08379298448562622
step: 120, loss: 0.019660141319036484
step: 130, loss: 3.9611521060578525e-05
step: 140, loss: 2.5457435185671784e-05
step: 150, loss: 9.22822582651861e-05
step: 160, loss: 0.0007233660435304046
step: 170, loss: 0.0004537104396149516
step: 180, loss: 1.8894343156716786e-05
step: 190, loss: 1.94600761460606e-05
step: 200, loss: 8.143021841533482e-05
step: 210, loss: 1.832779344113078e-05
step: 220, loss: 0.0019472470739856362
step: 230, loss: 0.00022770265059079975
/usr/local/lib/python3.7/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=0.9899441340782122, f1=0.9800443458980044,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.8864620869862847e-05
step: 10, loss: 0.013070477172732353
step: 20, loss: 4.051226278534159e-05
step: 30, loss: 0.00022178611834533513
step: 40, loss: 0.014382904395461082
step: 50, loss: 2.3438702555722557e-05
step: 60, loss: 1.6994556062854826e-05
step: 70, loss: 9.520756429992616e-05
step: 80, loss: 2.6955141947837546e-05
step: 90, loss: 3.1320538255386055e-05
step: 100, loss: 1.5921725207590498e-05
step: 110, loss: 0.0007115916814655066
step: 120, loss: 1.380952016916126e-05
step: 130, loss: 2.1419884433271363e-05
step: 140, loss: 4.1385050280950963e-05
step: 150, loss: 4.327265196479857e-05
step: 160, loss: 0.00665491446852684
step: 170, loss: 3.8092450267868116e-05
step: 180, loss: 6.947840302018449e-05
step: 190, loss: 5.200034865993075e-05
step: 200, loss: 2.0544242943287827e-05
step: 210, loss: 3.9781556552043185e-05
step: 220, loss: 0.00014304638898465782
step: 230, loss: 4.616933438228443e-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=0.9910313901345291, f1=0.9811320754716982,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.2390256819780916e-05
step: 10, loss: 1.2993713426112663e-05
step: 20, loss: 2.5502456992398947e-05
step: 30, loss: 3.776546873268671e-05
step: 40, loss: 0.00029586657183244824
step: 50, loss: 0.045393846929073334
step: 60, loss: 2.629458140290808e-05
step: 70, loss: 4.178529707132839e-05
step: 80, loss: 0.00027287553530186415

```

```

step: 90, loss: 0.014291934669017792
step: 100, loss: 3.730361277121119e-05
step: 110, loss: 1.6022275303839706e-05
step: 120, loss: 8.65996116772294e-05
step: 130, loss: 0.018886687234044075
step: 140, loss: 2.7975400371360593e-05
step: 150, loss: 6.625131936743855e-05
step: 160, loss: 7.753074896754697e-05
step: 170, loss: 1.6752299416111782e-05
step: 180, loss: 0.00023584645532537252
step: 190, loss: 0.0005023582489229739
step: 200, loss: 3.123118949588388e-05
step: 210, loss: 5.361914372770116e-05
step: 220, loss: 6.675948679912835e-05
step: 230, loss: 1.4297545931185596e-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=0.9910313901345291, f1=0.9811320754716982,
best_f1=0.9886877828054299
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:13, 181.77it/s]
load_f1 = 0.9943630214205187
real_f1 = 0.9932432432432432
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 235.68it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```


Downloading: 100% 433/433 [00:00<00:00, 417kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.62MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 69.9MB/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: 0.6133646368980408
step: 10, loss: 0.5268386602401733
step: 20, loss: 0.4306756556034088
step: 30, loss: 0.06702349334955215
step: 40, loss: 0.12762895226478577
step: 50, loss: 0.19574680924415588
step: 60, loss: 0.11087748408317566
step: 70, loss: 0.10280785709619522
step: 80, loss: 0.05214695259928703

```
step: 90, loss: 0.0673823431134224
step: 100, loss: 0.015144074335694313
step: 110, loss: 0.08248597383499146
step: 120, loss: 0.05729737877845764
step: 130, loss: 0.12147265672683716
step: 140, loss: 0.04804702475667
step: 150, loss: 0.021139590069651604
step: 160, loss: 0.02426670305430889
step: 170, loss: 0.15904779732227325
step: 180, loss: 0.08161719143390656
step: 190, loss: 0.00941552221775055
step: 200, loss: 0.15237924456596375
step: 210, loss: 0.08615104109048843
step: 220, loss: 0.22380812466144562
step: 230, loss: 0.1582520306110382
step: 240, loss: 0.04702095314860344
step: 250, loss: 0.02756255865097046
step: 260, loss: 0.05126703158020973
step: 270, loss: 0.011149166151881218
step: 280, loss: 0.022072862833738327
step: 290, loss: 0.09190501272678375
step: 300, loss: 0.014304403215646744
step: 310, loss: 0.18179373443126678
step: 320, loss: 0.044806934893131256
step: 330, loss: 0.010941467247903347
step: 340, loss: 0.06683346629142761
step: 350, loss: 0.14202992618083954
step: 360, loss: 0.032216522842645645
step: 370, loss: 0.0965527892112732
step: 380, loss: 0.02897597663104534
step: 390, loss: 0.09013836085796356
step: 400, loss: 0.21079710125923157
step: 410, loss: 0.04562757536768913
step: 420, loss: 0.010918055661022663
step: 430, loss: 0.1692364662885666
step: 440, loss: 0.016901127994060516
step: 450, loss: 0.001652448670938611
step: 460, loss: 0.010643196292221546
step: 470, loss: 0.1883886754512787
step: 480, loss: 0.040986936539411545
step: 490, loss: 0.16720430552959442
step: 500, loss: 0.08335140347480774
step: 510, loss: 0.010566785000264645
step: 520, loss: 0.026354722678661346
step: 530, loss: 0.0024156421422958374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9411210551106923, f1=0.9421720733427363,
best_f1=0.9421720733427363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14408239722251892
step: 10, loss: 0.028471358120441437
step: 20, loss: 0.02721838839352131
step: 30, loss: 0.03672987222671509
step: 40, loss: 0.0868128165602684
step: 50, loss: 0.055471476167440414
step: 60, loss: 0.06274694949388504
step: 70, loss: 0.035706616938114166
step: 80, loss: 0.0207215603441
step: 90, loss: 0.015389873646199703
step: 100, loss: 0.02150852605700493
step: 110, loss: 0.008564132265746593
step: 120, loss: 0.015480948612093925
step: 130, loss: 0.06415660679340363
step: 140, loss: 0.08110523968935013
step: 150, loss: 0.10145842283964157
step: 160, loss: 0.020569348707795143
step: 170, loss: 0.01394167821854353
step: 180, loss: 0.04622264578938484
step: 190, loss: 0.03312552347779274
step: 200, loss: 0.0047777676954865456
step: 210, loss: 0.0642315223813057
step: 220, loss: 0.04814204201102257
step: 230, loss: 0.006625866983085871
step: 240, loss: 0.062372539192438126
step: 250, loss: 0.03419140726327896
step: 260, loss: 0.008951866999268532
step: 270, loss: 0.0722779631614685
step: 280, loss: 0.01669531688094139
step: 290, loss: 0.03755725547671318
step: 300, loss: 0.15500622987747192
step: 310, loss: 0.01890253648161888
step: 320, loss: 0.09234003722667694
step: 330, loss: 0.033062394708395004
step: 340, loss: 0.01287865824997425
step: 350, loss: 0.0009812714997678995
step: 360, loss: 0.0965760126709938
step: 370, loss: 0.09952034801244736
step: 380, loss: 0.03516755625605583

```

```

step: 390, loss: 0.032365623861551285
step: 400, loss: 0.00594960106536746
step: 410, loss: 0.02856709249317646
step: 420, loss: 0.05459859222173691
step: 430, loss: 0.0294051431119442
step: 440, loss: 0.13948126137256622
step: 450, loss: 0.04171205684542656
step: 460, loss: 0.06203310564160347
step: 470, loss: 0.058681420981884
step: 480, loss: 0.24405933916568756
step: 490, loss: 0.01617036759853363
step: 500, loss: 0.18101340532302856
step: 510, loss: 0.013307640329003334
step: 520, loss: 0.04679805040359497
step: 530, loss: 0.006636898498982191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9453053783044667, f1=0.9409627611262488,
best_f1=0.9409627611262488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05383498966693878
step: 10, loss: 0.037334855645895004
step: 20, loss: 0.012445894069969654
step: 30, loss: 0.047943517565727234
step: 40, loss: 0.002446271013468504
step: 50, loss: 0.15862366557121277
step: 60, loss: 0.09266915172338486
step: 70, loss: 0.02243613265454769
step: 80, loss: 0.0052781300619244576
step: 90, loss: 0.006454831920564175
step: 100, loss: 0.03628566116094589
step: 110, loss: 0.000620781909674406
step: 120, loss: 0.006563093978911638
step: 130, loss: 0.0016345060430467129
step: 140, loss: 0.03184705227613449
step: 150, loss: 0.021711032837629318
step: 160, loss: 0.010667973197996616
step: 170, loss: 0.024701295420527458
step: 180, loss: 0.034546054899692535
step: 190, loss: 0.0031886021606624126
step: 200, loss: 0.0701230987906456

```

```

step: 210, loss: 0.08103708922863007
step: 220, loss: 0.03748379275202751
step: 230, loss: 0.08731389790773392
step: 240, loss: 0.004800578113645315
step: 250, loss: 0.03810126706957817
step: 260, loss: 0.019056398421525955
step: 270, loss: 0.01248146966099739
step: 280, loss: 0.214656263589859
step: 290, loss: 0.003519180929288268
step: 300, loss: 0.051783207803964615
step: 310, loss: 0.015760334208607674
step: 320, loss: 0.013851798139512539
step: 330, loss: 0.004969727713614702
step: 340, loss: 0.003249154891818762
step: 350, loss: 0.0165423434227705
step: 360, loss: 0.01988333649933338
step: 370, loss: 0.005617237649857998
step: 380, loss: 0.025417501106858253
step: 390, loss: 0.01126922108232975
step: 400, loss: 0.02153829298913479
step: 410, loss: 0.010839156806468964
step: 420, loss: 0.20609375834465027
step: 430, loss: 0.009464774280786514
step: 440, loss: 0.013024996966123581
step: 450, loss: 0.0823092982172966
step: 460, loss: 0.07009189575910568
step: 470, loss: 0.10394801944494247
step: 480, loss: 0.006139917764812708
step: 490, loss: 0.010117482393980026
step: 500, loss: 0.004122740123420954
step: 510, loss: 0.015557677485048771
step: 520, loss: 0.07833456993103027
step: 530, loss: 0.0228047464042902
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9505597014925373, f1=0.9490268767377201,
best_f1=0.9490268767377201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003622113261371851
step: 10, loss: 0.0027583749033510685
step: 20, loss: 0.0007032919675111771

```

step: 30, loss: 0.0009910777444019914
step: 40, loss: 0.004559554625302553
step: 50, loss: 0.04713818430900574
step: 60, loss: 0.00246093119494617
step: 70, loss: 0.06981941312551498
step: 80, loss: 0.01735260710120201
step: 90, loss: 0.012616511434316635
step: 100, loss: 0.0034108851104974747
step: 110, loss: 0.049523886293172836
step: 120, loss: 0.0057784151285886765
step: 130, loss: 0.005637339316308498
step: 140, loss: 0.012518368661403656
step: 150, loss: 0.009552329778671265
step: 160, loss: 0.0018137663137167692
step: 170, loss: 0.008715367875993252
step: 180, loss: 0.0036531100049614906
step: 190, loss: 0.3099323809146881
step: 200, loss: 0.01283945981413126
step: 210, loss: 0.015468595549464226
step: 220, loss: 0.005219207610934973
step: 230, loss: 0.07969848066568375
step: 240, loss: 0.0037098934408277273
step: 250, loss: 0.002873091259971261
step: 260, loss: 0.00241839955560863
step: 270, loss: 0.00455248961225152
step: 280, loss: 0.009227036498486996
step: 290, loss: 0.015041898004710674
step: 300, loss: 0.00018076063133776188
step: 310, loss: 0.025052864104509354
step: 320, loss: 0.0038686792831867933
step: 330, loss: 0.007039001677185297
step: 340, loss: 0.09004256129264832
step: 350, loss: 0.013009462505578995
step: 360, loss: 0.024951834231615067
step: 370, loss: 0.00267926137894392
step: 380, loss: 0.02207745984196663
step: 390, loss: 0.0039171879179775715
step: 400, loss: 0.03961533308029175
step: 410, loss: 0.001840354292653501
step: 420, loss: 0.1322813630104065
step: 430, loss: 0.02168220281600952
step: 440, loss: 0.010523626580834389
step: 450, loss: 0.007747157942503691
step: 460, loss: 0.0036382656544446945
step: 470, loss: 0.00893678329885006
step: 480, loss: 0.14148476719856262
step: 490, loss: 0.003851919434964657
step: 500, loss: 0.006626238580793142

```

step: 510, loss: 0.06608880311250687
step: 520, loss: 0.0349658839404583
step: 530, loss: 0.024630695581436157
/usr/local/lib/python3.7/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=0.9402501157943491, f1=0.9405255878284925,
best_f1=0.9490268767377201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01620258204638958
step: 10, loss: 0.03719215467572212
step: 20, loss: 0.004329627845436335
step: 30, loss: 0.001934210886247456
step: 40, loss: 0.034281857311725616
step: 50, loss: 0.003174091689288616
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 60, loss: 0.0064916666597127914
step: 70, loss: 0.00324477581307292
step: 80, loss: 0.0009682931704446673
step: 90, loss: 0.08574491739273071
step: 100, loss: 0.12562036514282227
step: 110, loss: 0.001187091344036162
step: 120, loss: 0.0031093826983124018
step: 130, loss: 0.001863003009930253
step: 140, loss: 0.005219544284045696
step: 150, loss: 0.014791049994528294
step: 160, loss: 0.000449452199973166
step: 170, loss: 0.03714264929294586
step: 180, loss: 0.0007316896226257086
step: 190, loss: 0.00048641394823789597
step: 200, loss: 0.0017803804948925972
step: 210, loss: 0.0036891603376716375
step: 220, loss: 0.04546350613236427
step: 230, loss: 0.007379340473562479
step: 240, loss: 0.005758793093264103
step: 250, loss: 0.0009480529115535319
step: 260, loss: 0.020425353199243546
step: 270, loss: 0.0018609168473631144
step: 280, loss: 0.018501223996281624
step: 290, loss: 0.12812450528144836
step: 300, loss: 0.006362674757838249
step: 310, loss: 0.004740845877677202
step: 320, loss: 0.0902051106095314
step: 330, loss: 0.05893510580062866
step: 340, loss: 0.0038259911816567183
step: 350, loss: 0.005270626395940781

```

```
step: 360, loss: 0.025604436174035072
step: 370, loss: 0.015654409304261208
step: 380, loss: 0.0002340872451895848
step: 390, loss: 0.000207549863262102
step: 400, loss: 0.003011573338881135
step: 410, loss: 0.011845833621919155
step: 420, loss: 0.00029756000731140375
step: 430, loss: 0.0013086777180433273
step: 440, loss: 0.053515270352363586
step: 450, loss: 0.0007925771060399711
step: 460, loss: 0.00027193553978577256
step: 470, loss: 0.006387244910001755
step: 480, loss: 0.04389965906739235
step: 490, loss: 0.002027959330007434
step: 500, loss: 0.011981809511780739
step: 510, loss: 0.009985500946640968
step: 520, loss: 0.005043565295636654
step: 530, loss: 0.033686909824609756
/usr/local/lib/python3.7/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=0.9480519480519481, f1=0.9424758398527383,
best_f1=0.9490268767377201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002586434595286846
step: 10, loss: 0.09902004152536392
step: 20, loss: 0.07263118773698807
step: 30, loss: 0.0057158442214131355
step: 40, loss: 0.027418509125709534
step: 50, loss: 0.03507227078080177
step: 60, loss: 0.00018467720656190068
step: 70, loss: 0.00028265698347240686
step: 80, loss: 0.06670846790075302
step: 90, loss: 0.0005416963831521571
step: 100, loss: 0.0007668787729926407
step: 110, loss: 0.0033630437683314085
step: 120, loss: 0.002248455071821809
step: 130, loss: 0.00020891909662168473
step: 140, loss: 0.020591503009200096
step: 150, loss: 0.0031462006736546755
step: 160, loss: 0.0007664369768463075
step: 170, loss: 0.0001115282138925977
step: 180, loss: 0.009190251119434834
step: 190, loss: 0.0023241047747433186
step: 200, loss: 0.00840581301599741
step: 210, loss: 0.006820052862167358
```



```

step: 220, loss: 0.009702137671411037
step: 230, loss: 0.005902417935431004
step: 240, loss: 0.08088890463113785
step: 250, loss: 0.16265612840652466
step: 260, loss: 0.0007677815738134086
step: 270, loss: 0.10231292992830276
step: 280, loss: 0.012889930047094822
step: 290, loss: 0.000977962277829647
step: 300, loss: 0.011373470537364483
step: 310, loss: 0.0019948354456573725
step: 320, loss: 0.0009611003915779293
step: 330, loss: 0.0001674590166658163
step: 340, loss: 0.06785929203033447
step: 350, loss: 0.007454876787960529
step: 360, loss: 0.0279671810567379
step: 370, loss: 0.022993816062808037
step: 380, loss: 0.0008379901410080492
step: 390, loss: 0.009838968515396118
step: 400, loss: 7.504875247832388e-05
step: 410, loss: 0.005376197397708893
step: 420, loss: 0.012775108218193054
step: 430, loss: 0.00028493086574599147
step: 440, loss: 0.00010895916784647852
step: 450, loss: 7.069246930768713e-05
step: 460, loss: 0.0002901200205087662
step: 470, loss: 0.0022039199247956276
step: 480, loss: 0.013470101170241833
step: 490, loss: 0.009939530864357948
step: 500, loss: 0.001517774537205696
step: 510, loss: 0.006148890592157841
step: 520, loss: 0.0033885191660374403
step: 530, loss: 0.030366996303200722
/usr/local/lib/python3.7/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=0.9490124023886082, f1=0.9484253765403924,
best_f1=0.9490268767377201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044876743108034134
step: 10, loss: 0.002918010810390115
step: 20, loss: 0.024076052010059357
step: 30, loss: 0.001141197164542973
step: 40, loss: 0.0010576702188700438
step: 50, loss: 0.0325382724404335
step: 60, loss: 0.0018086563795804977
step: 70, loss: 0.008224924094974995

```

step: 80, loss: 0.002245877403765917
step: 90, loss: 0.0017948879394680262
step: 100, loss: 0.00011668524530250579
step: 110, loss: 0.0003951344988308847
step: 120, loss: 0.019527338445186615
step: 130, loss: 0.0100325345993042
step: 140, loss: 0.0007881029741838574
step: 150, loss: 0.00020123919239267707
step: 160, loss: 7.486202957807109e-05
step: 170, loss: 0.00012775510549545288
step: 180, loss: 0.001086209318600595
step: 190, loss: 0.001059921458363533
step: 200, loss: 0.00046574673615396023
step: 210, loss: 0.055412907153367996
step: 220, loss: 0.026461509987711906
step: 230, loss: 0.034033242613077164
step: 240, loss: 0.008742572739720345
step: 250, loss: 0.029781658202409744
step: 260, loss: 0.0011771861463785172
step: 270, loss: 0.10856201499700546
step: 280, loss: 0.014523182064294815
step: 290, loss: 0.0004079495556652546
step: 300, loss: 0.0044419774785637856
step: 310, loss: 5.8014262322103605e-05
step: 320, loss: 0.00016950542340055108
step: 330, loss: 0.02400633506476879
step: 340, loss: 0.03103860281407833
step: 350, loss: 0.0013778676511719823
step: 360, loss: 0.00770470779389143
step: 370, loss: 0.006555704399943352
step: 380, loss: 0.0919199138879776
step: 390, loss: 0.00034601319930516183
step: 400, loss: 0.02909017540514469
step: 410, loss: 0.0016939707566052675
step: 420, loss: 0.0003698837535921484
step: 430, loss: 8.305055962409824e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.13782982528209686
step: 450, loss: 5.271990448818542e-05
step: 460, loss: 0.0011866958811879158
step: 470, loss: 0.15007805824279785
step: 480, loss: 0.05266387015581131
step: 490, loss: 0.0012111137621104717
step: 500, loss: 0.00039734120946377516
step: 510, loss: 0.004739544820040464
step: 520, loss: 0.024251410737633705
step: 530, loss: 0.027454804629087448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9544186046511628, f1=0.9489185457892315,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018910422921180725
step: 10, loss: 0.06084943562746048
step: 20, loss: 9.24504129216075e-05
step: 30, loss: 0.015831585973501205
step: 40, loss: 9.511333337286487e-05
step: 50, loss: 0.0035676993429660797
step: 60, loss: 0.0011603590101003647
step: 70, loss: 0.0008662217878736556
step: 80, loss: 0.0010727044427767396
step: 90, loss: 0.0014340556226670742
step: 100, loss: 0.0002544762974139303
step: 110, loss: 0.00016345940821338445
step: 120, loss: 0.0019243273418396711
step: 130, loss: 0.0008209627121686935
step: 140, loss: 0.0075945062562823296
step: 150, loss: 0.0075331865809857845
step: 160, loss: 0.0184926800429821
step: 170, loss: 0.06991976499557495
step: 180, loss: 0.000745475641451776
step: 190, loss: 0.0033009075559675694
step: 200, loss: 0.0025233356282114983
step: 210, loss: 0.01776430942118168
step: 220, loss: 0.003660304704681039
step: 230, loss: 0.00029433914460241795
step: 240, loss: 0.0020189080387353897
step: 250, loss: 0.0001841978810261935
step: 260, loss: 0.000160043899086304
step: 270, loss: 0.00661154231056571
step: 280, loss: 0.017375024035573006
step: 290, loss: 0.0009983985219150782
step: 300, loss: 0.0018338526133447886
step: 310, loss: 0.01155297551304102
step: 320, loss: 0.012462827377021313
step: 330, loss: 0.0029368263203650713
step: 340, loss: 0.0002098119875881821
step: 350, loss: 0.00025605782866477966
step: 360, loss: 0.016698485240340233

```

```
step: 370, loss: 0.00012790433538611978
step: 380, loss: 0.002978523727506399
step: 390, loss: 0.13941460847854614
step: 400, loss: 0.0010624597780406475
step: 410, loss: 0.0004357906000223011
step: 420, loss: 0.001736003439873457
step: 430, loss: 0.05463080480694771
step: 440, loss: 0.003420011606067419
step: 450, loss: 0.0010398589074611664
step: 460, loss: 0.012365917675197124
step: 470, loss: 0.0015748862642794847
step: 480, loss: 0.03708763048052788
step: 490, loss: 0.001177144586108625
step: 500, loss: 0.00038779000169597566
step: 510, loss: 0.0033125276677310467
step: 520, loss: 0.001456660102121532
step: 530, loss: 0.001139480504207313
/usr/local/lib/python3.7/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=0.9463781749764816, f1=0.9515828677839852,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027012373320758343
step: 10, loss: 0.09863199293613434
step: 20, loss: 0.003510685870423913
step: 30, loss: 0.0006453492678701878
step: 40, loss: 0.08493898063898087
step: 50, loss: 0.0012547837104648352
step: 60, loss: 0.0002849537122529
step: 70, loss: 0.03913377225399017
step: 80, loss: 0.003152046585455537
step: 90, loss: 7.259649282786995e-05
step: 100, loss: 0.0001228220935445279
step: 110, loss: 0.00013470165140461177
step: 120, loss: 0.0011990992352366447
step: 130, loss: 0.002843710361048579
step: 140, loss: 0.01171198021620512
step: 150, loss: 0.00029209692729637027
step: 160, loss: 0.001495881355367601
step: 170, loss: 0.06977440416812897
step: 180, loss: 0.0005276093725115061
step: 190, loss: 0.00016177516954485327
step: 200, loss: 0.00036359051591716707
step: 210, loss: 0.000510913843754679
step: 220, loss: 0.0009403208387084305
```

```

step: 230, loss: 0.00015819413238205016
step: 240, loss: 0.0011386264814063907
step: 250, loss: 0.00017606018809601665
step: 260, loss: 0.00970964040607214
step: 270, loss: 0.00013652608322445303
step: 280, loss: 0.005549861583858728
step: 290, loss: 0.03800082579255104
step: 300, loss: 5.6688455515541136e-05
step: 310, loss: 0.023493262007832527
step: 320, loss: 0.00414310023188591
step: 330, loss: 0.0002585187612567097
step: 340, loss: 3.499874583212659e-05
step: 350, loss: 0.008249813690781593
step: 360, loss: 0.001552075962536037
step: 370, loss: 0.002216675551608205
step: 380, loss: 0.0001407653617206961
step: 390, loss: 0.0006382198189385235
step: 400, loss: 0.00038922211388126016
step: 410, loss: 9.318601223640144e-05
step: 420, loss: 2.4980985472211614e-05
step: 430, loss: 4.50249899586197e-05
step: 440, loss: 0.0020646119955927134
step: 450, loss: 7.039852062007412e-05
step: 460, loss: 0.00010190895409323275
step: 470, loss: 2.570716242189519e-05
step: 480, loss: 0.0014696450671181083
step: 490, loss: 0.013817977160215378
step: 500, loss: 0.0019088606350123882
step: 510, loss: 0.0024497313424944878
step: 520, loss: 0.036707803606987
step: 530, loss: 3.397691762074828e-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=0.9450549450549451, f1=0.9449749886311961,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002634330652654171
step: 10, loss: 2.591963675513398e-05
step: 20, loss: 3.963890776503831e-05
step: 30, loss: 0.00025193265173584223
step: 40, loss: 5.801173028885387e-05
step: 50, loss: 0.000721563643310219
step: 60, loss: 5.985501047689468e-05
step: 70, loss: 0.0005860555684193969
step: 80, loss: 4.631202682503499e-05

```

```
step: 90, loss: 0.0009435409447178245
step: 100, loss: 0.00016231063636951149
step: 110, loss: 3.6064480809727684e-05
step: 120, loss: 0.00012049451470375061
step: 130, loss: 3.711085446411744e-05
step: 140, loss: 0.016086922958493233
step: 150, loss: 0.005740530788898468
step: 160, loss: 3.11255753331352e-05
step: 170, loss: 0.0005372504820115864
step: 180, loss: 0.021369898691773415
step: 190, loss: 0.0001168398666777648
step: 200, loss: 3.1554132874589413e-05
step: 210, loss: 0.01952766627073288
step: 220, loss: 0.006707658059895039
step: 230, loss: 0.0021099229343235493
step: 240, loss: 0.0010086894035339355
step: 250, loss: 3.8280551962088794e-05
step: 260, loss: 0.0013010582188144326
step: 270, loss: 0.0006014542886987329
step: 280, loss: 3.078384543186985e-05
step: 290, loss: 8.507470192853361e-05
step: 300, loss: 1.892777618195396e-05
step: 310, loss: 6.500577001133934e-05
step: 320, loss: 0.0018052097875624895
step: 330, loss: 0.00017628150817472488
step: 340, loss: 0.0005916391382925212
step: 350, loss: 0.0008441780228167772
step: 360, loss: 0.000786570250056684
step: 370, loss: 0.0015772731276229024
step: 380, loss: 0.03263282775878906
step: 390, loss: 5.989014607621357e-05
step: 400, loss: 0.0023890873417258263
step: 410, loss: 0.14869816601276398
step: 420, loss: 0.0004709597851615399
step: 430, loss: 9.07186622498557e-05
step: 440, loss: 0.006110093556344509
step: 450, loss: 5.703365241060965e-05
step: 460, loss: 0.0033702775835990906
step: 470, loss: 0.001753311138600111
step: 480, loss: 0.0002370672591496259
step: 490, loss: 0.00010728406778071076
step: 500, loss: 0.008589676581323147
step: 510, loss: 0.0007799494778737426
step: 520, loss: 0.0008100529666990042
step: 530, loss: 0.0021323058754205704
/usr/local/lib/python3.7/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=0.9474662947466295, f1=0.9529085872576176,  
best_f1=0.9489185457892315  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.000549479213077575  
step: 10, loss: 0.004179461393505335  
step: 20, loss: 5.92392279941123e-05  
step: 30, loss: 3.662494782474823e-05  
step: 40, loss: 0.0005065599107183516  
step: 50, loss: 0.00202256441116333  
step: 60, loss: 0.02076859027147293  
step: 70, loss: 0.00026296722353436053  
step: 80, loss: 0.0012356775114312768  
step: 90, loss: 1.9057959434576333e-05  
step: 100, loss: 0.007585730869323015  
step: 110, loss: 0.000435314723290503  
step: 120, loss: 0.0001267252810066566  
step: 130, loss: 5.1166418415959924e-05  
step: 140, loss: 0.004578732885420322  
step: 150, loss: 3.273982656537555e-05  
step: 160, loss: 9.17783472687006e-05  
step: 170, loss: 0.0016826116479933262  
step: 180, loss: 0.00042786187259480357  
step: 190, loss: 0.0046832021325826645  
step: 200, loss: 0.00708577036857605  
step: 210, loss: 0.0003802580467890948  
step: 220, loss: 0.0020747326780110598  
step: 230, loss: 0.000622407766059041  
step: 240, loss: 0.002049849834293127  
step: 250, loss: 0.00017694330017548054  
step: 260, loss: 0.00025152022135443985  
step: 270, loss: 0.04366638511419296  
step: 280, loss: 0.0008492969791404903  
step: 290, loss: 0.0002381087397225201  
step: 300, loss: 0.0007355199195444584  
step: 310, loss: 0.00021739656222052872  
step: 320, loss: 7.570943125756457e-05  
step: 330, loss: 0.0003367481112945825  
step: 340, loss: 7.936568727018312e-05  
step: 350, loss: 0.0005324435187503695  
step: 360, loss: 0.00011300744517939165  
step: 370, loss: 0.003916414920240641  
step: 380, loss: 0.0003845398314297199  
step: 390, loss: 7.215538789751008e-05  
step: 400, loss: 0.000367264001397416  
step: 410, loss: 0.0006874519749544561  
step: 420, loss: 0.0005069856415502727
```

```

step: 430, loss: 8.911181066650897e-05
step: 440, loss: 0.000749509665183723
step: 450, loss: 0.0008667460060678422
step: 460, loss: 0.2026824653148651
step: 470, loss: 0.00027005517040379345
step: 480, loss: 0.00239067361690104
step: 490, loss: 0.000872405304107815
step: 500, loss: 0.00015909005014691502
step: 510, loss: 0.007089503575116396
step: 520, loss: 3.180301791871898e-05
step: 530, loss: 0.00026509276358410716
/usr/local/lib/python3.7/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=0.94682352941117646, f1=0.9492787342950209,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010647304588928819
step: 10, loss: 4.754757173941471e-05
step: 20, loss: 7.815566641511396e-05
step: 30, loss: 0.0021748440340161324
step: 40, loss: 0.0005140642751939595
step: 50, loss: 0.04951760545372963
step: 60, loss: 0.0019645157735794783
step: 70, loss: 0.0016453908756375313
step: 80, loss: 0.00014254609413910657
step: 90, loss: 0.00014225783525034785
step: 100, loss: 7.70328551880084e-05
step: 110, loss: 8.833380707073957e-05
step: 120, loss: 0.00018402257410343736
step: 130, loss: 0.001841556979343295
step: 140, loss: 5.7505956647219136e-05
step: 150, loss: 5.748541298089549e-05
step: 160, loss: 8.968496695160866e-05
step: 170, loss: 6.952572584850714e-05
step: 180, loss: 9.247818525182083e-05
step: 190, loss: 0.00046622566878795624
step: 200, loss: 0.00020623212913051248
step: 210, loss: 0.00019190419698134065
step: 220, loss: 0.00013511358702089638
step: 230, loss: 5.6145603593904525e-05
step: 240, loss: 0.0002947275643236935
step: 250, loss: 7.34187924535945e-05
step: 260, loss: 0.0004728573840111494
step: 270, loss: 0.007409048732370138
step: 280, loss: 0.0008397759520448744

```



```

step: 290, loss: 0.00024710394791327417
step: 300, loss: 0.009232074953615665
step: 310, loss: 5.1165490731364116e-05
step: 320, loss: 1.6659156244713813e-05
step: 330, loss: 0.007949761115014553
step: 340, loss: 0.0003030011139344424
step: 350, loss: 0.00016850882093422115
step: 360, loss: 0.0382915697991848
step: 370, loss: 0.00961021427065134
step: 380, loss: 0.00015241299115587026
step: 390, loss: 5.714234430342913e-05
step: 400, loss: 0.0016339439898729324
step: 410, loss: 0.009253805503249168
step: 420, loss: 0.006309698801487684
step: 430, loss: 0.01525876298546791
step: 440, loss: 0.0016415388090535998
step: 450, loss: 0.0005102814175188541
step: 460, loss: 0.00024271567235700786
step: 470, loss: 0.0016713814111426473
step: 480, loss: 0.0024001230485737324
step: 490, loss: 0.0009074881090782583
step: 500, loss: 0.0002437368966639042
step: 510, loss: 0.0013768678763881326
step: 520, loss: 0.00016040496120695025
step: 530, loss: 0.0017625398468226194
/usr/local/lib/python3.7/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=0.9516728624535317, f1=0.9469801751959428,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005780500825494528
step: 10, loss: 4.949812500854023e-05
step: 20, loss: 0.01762317679822445
step: 30, loss: 0.00012920501467306167
step: 40, loss: 0.005608972627669573
step: 50, loss: 0.010318964719772339
step: 60, loss: 0.004529422614723444
step: 70, loss: 8.015828643692657e-05
step: 80, loss: 4.339336373959668e-05
step: 90, loss: 0.030593037605285645
step: 100, loss: 0.0032723008189350367
step: 110, loss: 7.424153591273353e-05
step: 120, loss: 0.00019840881577692926
step: 130, loss: 4.801448085345328e-05
step: 140, loss: 3.2609703339403495e-05

```

```

step: 150, loss: 6.188283441588283e-05
step: 160, loss: 0.0007039067568257451
step: 170, loss: 0.004525774158537388
step: 180, loss: 3.685513001983054e-05
step: 190, loss: 0.00038005379610694945
step: 200, loss: 2.4819706595735624e-05
step: 210, loss: 5.228230656939559e-05
step: 220, loss: 3.876829214277677e-05
step: 230, loss: 9.283991676056758e-05
step: 240, loss: 3.62180799129419e-05
step: 250, loss: 4.202622585580684e-05
step: 260, loss: 0.0020155960228294134
step: 270, loss: 2.0008119463454932e-05
step: 280, loss: 5.1105198508594185e-05
step: 290, loss: 0.001101725734770298
step: 300, loss: 3.348489917698316e-05
step: 310, loss: 0.000988891813904047
step: 320, loss: 0.0026461011730134487
step: 330, loss: 0.0008115273085422814
step: 340, loss: 0.0003276122733950615
step: 350, loss: 0.00020123088324908167
step: 360, loss: 2.7498239433043636e-05
step: 370, loss: 0.00013340445002540946
step: 380, loss: 2.3721933757769875e-05
step: 390, loss: 0.0008031336474232376
step: 400, loss: 0.00010199328971793875
step: 410, loss: 0.00268022739328444
step: 420, loss: 2.909635441028513e-05
step: 430, loss: 0.003073187777772546
step: 440, loss: 6.518948794109747e-05
step: 450, loss: 3.754587669391185e-05
step: 460, loss: 0.08561551570892334
step: 470, loss: 7.385612843791023e-05
step: 480, loss: 0.0004407072556205094
step: 490, loss: 0.0003680216905195266
step: 500, loss: 0.00019916061137337238
step: 510, loss: 5.590896398643963e-05
step: 520, loss: 0.005125723779201508
step: 530, loss: 4.936976620228961e-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=0.9511737089201877, f1=0.9458955223880597,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010805762140080333

```

step: 10, loss: 2.3893417164799757e-05
step: 20, loss: 0.0013914508745074272
step: 30, loss: 0.0002527291653677821
step: 40, loss: 0.0001590097526786849
step: 50, loss: 0.0007193055935204029
step: 60, loss: 1.6368619981221855e-05
step: 70, loss: 0.00021009152987971902
step: 80, loss: 1.8052416635327972e-05
step: 90, loss: 0.0002914028300438076
step: 100, loss: 0.06364089995622635
step: 110, loss: 0.00012794183567166328
step: 120, loss: 8.770942804403603e-05
step: 130, loss: 0.020206645131111145
step: 140, loss: 0.0007839177269488573
step: 150, loss: 6.666243280051276e-05
step: 160, loss: 0.0004838325548917055
step: 170, loss: 0.00025965701206587255
step: 180, loss: 0.00014737641322426498
step: 190, loss: 0.00041364060598425567
step: 200, loss: 0.014152281917631626
step: 210, loss: 2.8220583772053942e-05
step: 220, loss: 2.653812407515943e-05
step: 230, loss: 0.00345470174215734
step: 240, loss: 0.0011467351578176022
step: 250, loss: 9.89785767160356e-05
step: 260, loss: 0.0029264448676258326
step: 270, loss: 0.00011871738388435915
step: 280, loss: 2.020547617576085e-05
step: 290, loss: 0.0009399177506566048
step: 300, loss: 0.0029520918615162373
step: 310, loss: 0.02372395060956478
step: 320, loss: 0.001392581150867045
step: 330, loss: 0.00041313128895126283
step: 340, loss: 0.0011299226898699999
step: 350, loss: 0.00026231599622406065
step: 360, loss: 0.0017826651455834508
step: 370, loss: 0.00020143869915045798
step: 380, loss: 5.767084439867176e-05
step: 390, loss: 0.02729777805507183
step: 400, loss: 0.00016337132547050714
step: 410, loss: 1.3545031833928078e-05
step: 420, loss: 0.0025858015287667513
step: 430, loss: 4.088829882675782e-05
step: 440, loss: 0.011089048348367214
step: 450, loss: 6.174272857606411e-05
step: 460, loss: 0.0003478083817753941
step: 470, loss: 3.009051943081431e-05
step: 480, loss: 1.60296076501254e-05

```

step: 490, loss: 2.3986096493899822e-05
step: 500, loss: 1.7560778360348195e-05
step: 510, loss: 0.0006294813356362283
step: 520, loss: 0.0003604497469495982
step: 530, loss: 1.8030132196145132e-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=0.9506517690875234, f1=0.9500924214417744,
best_f1=0.9489185457892315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.053480708738789e-05
step: 10, loss: 2.620155100885313e-05
step: 20, loss: 0.00018414646910969168
step: 30, loss: 0.0003730652097146958
step: 40, loss: 0.09207790344953537
step: 50, loss: 0.009935147128999233
step: 60, loss: 3.242682942072861e-05
step: 70, loss: 0.000376175157725811
step: 80, loss: 8.77300844877027e-05
step: 90, loss: 1.9680350305861793e-05
step: 100, loss: 2.5848437871900387e-05
step: 110, loss: 0.0016068547265604138
step: 120, loss: 0.0003738352097570896
step: 130, loss: 0.1352892518043518
step: 140, loss: 4.855815495830029e-05
step: 150, loss: 4.348270158516243e-05
step: 160, loss: 5.1421800890238956e-05
step: 170, loss: 2.0525219952105545e-05
step: 180, loss: 0.00015300363884307444
step: 190, loss: 2.8414377084118314e-05
step: 200, loss: 2.6049827283713967e-05
step: 210, loss: 0.0016912902938202024
step: 220, loss: 8.822582458378747e-05
step: 230, loss: 0.00029149098554626107
step: 240, loss: 3.0191558835213073e-05
step: 250, loss: 5.638534275931306e-05
step: 260, loss: 0.00226496160030365
step: 270, loss: 3.439210559008643e-05
step: 280, loss: 1.544108090456575e-05
step: 290, loss: 1.4412981727218721e-05
step: 300, loss: 1.8823442587745376e-05
step: 310, loss: 9.086273348657414e-05
step: 320, loss: 0.002248827600851655
step: 330, loss: 0.00014176398690324277
step: 340, loss: 6.481683521997184e-05

```

```

step: 350, loss: 3.238946010242216e-05
step: 360, loss: 0.0017001634696498513
step: 370, loss: 0.0016341708833351731
step: 380, loss: 6.725010462105274e-05
step: 390, loss: 0.0035842079669237137
step: 400, loss: 6.492401735158637e-05
step: 410, loss: 9.579210745869204e-05
step: 420, loss: 9.901641897158697e-05
step: 430, loss: 0.0006445098551921546
step: 440, loss: 0.0009647021070122719
step: 450, loss: 4.8494570364709944e-05
step: 460, loss: 1.6968228010227904e-05
step: 470, loss: 0.0013223231071606278
step: 480, loss: 1.6204723578994162e-05
step: 490, loss: 1.1812779121100903e-05
step: 500, loss: 2.2048861865187064e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 510, loss: 0.0006033515091985464
step: 520, loss: 1.6513864466105588e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 530, loss: 0.000584998691920191
/usr/local/lib/python3.7/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=0.9515828677839852, f1=0.9491211840888066,
best_f1=0.9489185457892315
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:22, 252.58it/s]
load_f1 = 0.950622981079834
real_f1 = 0.950668510834486
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, 245.30it/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: 0.5692244172096252
step: 10, loss: 0.36565038561820984
step: 20, loss: 0.3580814003944397
step: 30, loss: 0.30252909660339355
step: 40, loss: 0.16955259442329407
```

```

step: 50, loss: 0.3200629651546478
step: 60, loss: 0.13015010952949524
step: 70, loss: 0.18341287970542908
step: 80, loss: 0.23749768733978271
step: 90, loss: 0.30559512972831726
step: 100, loss: 0.5131158828735352
step: 110, loss: 0.17635264992713928
step: 120, loss: 0.1985902637243271
step: 130, loss: 0.1039053276181221
step: 140, loss: 0.2407098263502121
step: 150, loss: 0.1526225209236145
step: 160, loss: 0.18081268668174744
step: 170, loss: 0.2015903890132904
step: 180, loss: 0.10587231069803238
step: 190, loss: 0.286810964345932
step: 200, loss: 0.20750536024570465
step: 210, loss: 0.23023095726966858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.693711967545639, f1=0.711018711018711,
best_f1=0.711018711018711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0799219012260437
step: 10, loss: 0.20007000863552094
step: 20, loss: 0.2931201756000519
step: 30, loss: 0.16707627475261688
step: 40, loss: 0.14429572224617004
step: 50, loss: 0.24022969603538513
step: 60, loss: 0.32824891805648804
step: 70, loss: 0.12401796877384186
step: 80, loss: 0.12523555755615234
step: 90, loss: 0.18940035998821259
step: 100, loss: 0.01605594903230667
step: 110, loss: 0.10615581274032593
step: 120, loss: 0.1833205223083496
step: 130, loss: 0.008718977682292461
step: 140, loss: 0.24435566365718842
step: 150, loss: 0.27640825510025024
step: 160, loss: 0.168046236038208
step: 170, loss: 0.13554643094539642
step: 180, loss: 0.16830234229564667

```



```

step: 190, loss: 0.16654188930988312
step: 200, loss: 0.07958647608757019
step: 210, loss: 0.16653190553188324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7330827067669173, f1=0.7153558052434457,
best_f1=0.7153558052434457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05993366986513138
step: 10, loss: 0.16729387640953064
step: 20, loss: 0.13893356919288635
step: 30, loss: 0.1987776756286621
step: 40, loss: 0.07345689833164215
step: 50, loss: 0.05465850606560707
step: 60, loss: 0.192353293299675
step: 70, loss: 0.0634976476430893
step: 80, loss: 0.23045586049556732
step: 90, loss: 0.05090408772230148
step: 100, loss: 0.1262853592634201
step: 110, loss: 0.08950388431549072
step: 120, loss: 0.17230454087257385
step: 130, loss: 0.21017000079154968
step: 140, loss: 0.05921398848295212
step: 150, loss: 0.29294586181640625
step: 160, loss: 0.009617346338927746
step: 170, loss: 0.09174007177352905
step: 180, loss: 0.1448388695716858
step: 190, loss: 0.2412736564874649
step: 200, loss: 0.008158058859407902
step: 210, loss: 0.10476057231426239
/usr/local/lib/python3.7/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=0.7193675889328063, f1=0.7306122448979591,
best_f1=0.7153558052434457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09396301209926605
step: 10, loss: 0.058306340128183365
step: 20, loss: 0.06020641699433327

```

```

step: 30, loss: 0.06694581359624863
step: 40, loss: 0.06291840970516205
step: 50, loss: 0.2068527638912201
step: 60, loss: 0.11783269047737122
step: 70, loss: 0.1417425274848938
step: 80, loss: 0.16909517347812653
step: 90, loss: 0.011344974860548973
step: 100, loss: 0.2630101442337036
step: 110, loss: 0.08964252471923828
step: 120, loss: 0.16985277831554413
step: 130, loss: 0.37211981415748596
step: 140, loss: 0.15524829924106598
step: 150, loss: 0.022818531841039658
step: 160, loss: 0.03368715941905975
step: 170, loss: 0.12196541577577591
step: 180, loss: 0.3682624399662018
step: 190, loss: 0.08528587967157364
step: 200, loss: 0.10975325107574463
step: 210, loss: 0.09719035774469376
/usr/local/lib/python3.7/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=0.7285714285714286, f1=0.7191413237924865,
best_f1=0.7153558052434457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09054655581712723
step: 10, loss: 0.032625146210193634
step: 20, loss: 0.2562299072742462
step: 30, loss: 0.08097442984580994
step: 40, loss: 0.0963435173034668
step: 50, loss: 0.06938263773918152
step: 60, loss: 0.10033156722784042
step: 70, loss: 0.10322225093841553
step: 80, loss: 0.03863795846700668
step: 90, loss: 0.14079773426055908
step: 100, loss: 0.03945547342300415
step: 110, loss: 0.17233780026435852
step: 120, loss: 0.23301483690738678
step: 130, loss: 0.06403909623622894
step: 140, loss: 0.0239765215665102
step: 150, loss: 0.06572864949703217
step: 160, loss: 0.16551271080970764
step: 170, loss: 0.07989273220300674
step: 180, loss: 0.10290100425481796
step: 190, loss: 0.022502394393086433
step: 200, loss: 0.0874113067984581

```

```

step: 210, loss: 0.029481537640094757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7387387387387387, f1=0.7380073800738007,
best_f1=0.7380073800738007
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03985924273729324
step: 10, loss: 0.03959081694483757
step: 20, loss: 0.03954426944255829
step: 30, loss: 0.00381935085169971
step: 40, loss: 0.02445465512573719
step: 50, loss: 0.005221846979111433
step: 60, loss: 0.15255174040794373
step: 70, loss: 0.06140679121017456
step: 80, loss: 0.3335344195365906
step: 90, loss: 0.14482128620147705
step: 100, loss: 0.03962050750851631
step: 110, loss: 0.028798090294003487
step: 120, loss: 0.129491925239563
step: 130, loss: 0.04756690561771393
step: 140, loss: 0.08401505649089813
step: 150, loss: 0.0628102719783783
step: 160, loss: 0.01105872355401516
step: 170, loss: 0.03215022012591362
step: 180, loss: 0.03168950974941254
step: 190, loss: 0.07839451730251312
step: 200, loss: 0.0036988933570683002
step: 210, loss: 0.04522893577814102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7486238532110092, f1=0.7193973634651601,
best_f1=0.7193973634651601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009606502950191498

```

```

step: 10, loss: 0.0043308306485414505
step: 20, loss: 0.045071374624967575
step: 30, loss: 0.10966141521930695
step: 40, loss: 0.04578367620706558
step: 50, loss: 0.07919227331876755
step: 60, loss: 0.07450885325670242
step: 70, loss: 0.007059271913021803
step: 80, loss: 0.10699725151062012
step: 90, loss: 0.026029985398054123
step: 100, loss: 0.018253345042467117
step: 110, loss: 0.22667858004570007
step: 120, loss: 0.3088268041610718
step: 130, loss: 0.04370672255754471
step: 140, loss: 0.015603910200297832
step: 150, loss: 0.007032616529613733
step: 160, loss: 0.1693502813577652
step: 170, loss: 0.09708572924137115
step: 180, loss: 0.02831723541021347
step: 190, loss: 0.1285737305879593
step: 200, loss: 0.0055671497248113155
step: 210, loss: 0.0940113514661789
/usr/local/lib/python3.7/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=0.736842105263158, f1=0.7140039447731756,
best_f1=0.7193973634651601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044126417487859726
step: 10, loss: 0.3089364469051361
step: 20, loss: 0.004761543590575457
step: 30, loss: 0.038412079215049744
step: 40, loss: 0.0019759717397391796
step: 50, loss: 0.020716369152069092
step: 60, loss: 0.21929048001766205
step: 70, loss: 0.06546112895011902
step: 80, loss: 0.027537876740098
step: 90, loss: 0.004374807234853506
step: 100, loss: 0.11883536726236343
step: 110, loss: 0.1463354527950287
step: 120, loss: 0.01807340979576111
step: 130, loss: 0.008450709283351898
step: 140, loss: 0.10705196112394333
step: 150, loss: 0.03829382359981537
step: 160, loss: 0.09786465764045715
step: 170, loss: 0.05721430107951164
step: 180, loss: 0.0919884741306305

```

```

step: 190, loss: 0.015382947400212288
step: 200, loss: 0.0030808548908680677
step: 210, loss: 0.27426108717918396
/usr/local/lib/python3.7/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=0.7462686567164181, f1=0.7251461988304092,
best_f1=0.7193973634651601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01598985306918621
step: 10, loss: 0.06565523147583008
step: 20, loss: 0.008553530089557171
step: 30, loss: 0.08343231678009033
step: 40, loss: 0.08577713370323181
step: 50, loss: 0.01722826436161995
step: 60, loss: 0.16299717128276825
step: 70, loss: 0.18178367614746094
step: 80, loss: 0.007983192801475525
step: 90, loss: 0.003910887520760298
step: 100, loss: 0.01255117915570736
step: 110, loss: 0.016996320337057114
step: 120, loss: 0.0634651631116867
step: 130, loss: 0.0030762634705752134
step: 140, loss: 0.05088839679956436
step: 150, loss: 0.14020618796348572
step: 160, loss: 0.0018094213446602225
step: 170, loss: 0.010989752598106861
step: 180, loss: 0.048884905874729156
step: 190, loss: 0.005732146091759205
step: 200, loss: 0.08400692045688629
step: 210, loss: 0.006723455619066954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7649253731343284, f1=0.7140115163147792,
best_f1=0.7140115163147792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005653888452798128
step: 10, loss: 0.026164792478084564
step: 20, loss: 0.0007819763268344104

```

```

step: 30, loss: 0.12559030950069427
step: 40, loss: 0.002623203909024596
step: 50, loss: 0.016724303364753723
step: 60, loss: 0.02588147670030594
step: 70, loss: 0.07258018106222153
step: 80, loss: 0.12163212895393372
step: 90, loss: 0.04197785630822182
step: 100, loss: 0.03377522900700569
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 110, loss: 0.003914844244718552
step: 120, loss: 0.014958023093640804
step: 130, loss: 0.14943577349185944
step: 140, loss: 0.020120741799473763
step: 150, loss: 0.0598708800971508
step: 160, loss: 0.0027674920856952667
step: 170, loss: 0.01446076575666666
step: 180, loss: 0.0908772200345993
step: 190, loss: 0.054105840623378754
step: 200, loss: 0.006790092680603266
step: 210, loss: 0.05079037323594093
/usr/local/lib/python3.7/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=0.7374517374517374, f1=0.7077534791252484,
best_f1=0.7140115163147792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036221373826265335
step: 10, loss: 0.0920819342136383
step: 20, loss: 0.07972324639558792
step: 30, loss: 0.000260250351857394
step: 40, loss: 0.04507846385240555
step: 50, loss: 0.007662687450647354
step: 60, loss: 0.02675606496632099
step: 70, loss: 0.013161438517272472
step: 80, loss: 0.2358683943748474
step: 90, loss: 0.0788917988538742
step: 100, loss: 0.02585136517882347
step: 110, loss: 0.027707507833838463
step: 120, loss: 0.070815309882164
step: 130, loss: 0.011541900224983692
step: 140, loss: 0.015446730889379978
step: 150, loss: 0.016916979104280472
step: 160, loss: 0.11366362124681473
step: 170, loss: 0.05435025319457054
step: 180, loss: 0.013423818163573742
step: 190, loss: 0.022915858775377274

```

```

step: 200, loss: 0.013135800138115883
step: 210, loss: 0.0006966134533286095
/usr/local/lib/python3.7/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=0.7454545454545455, f1=0.7012987012987012,
best_f1=0.7140115163147792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13864587247371674
step: 10, loss: 0.016786418855190277
step: 20, loss: 0.04280664771795273
step: 30, loss: 0.014073907397687435
step: 40, loss: 0.12085864692926407
step: 50, loss: 0.17722536623477936
step: 60, loss: 0.06771686673164368
step: 70, loss: 0.016404205933213234
step: 80, loss: 0.007148676551878452
step: 90, loss: 0.05379262566566467
step: 100, loss: 0.01829533651471138
step: 110, loss: 0.0038013376761227846
step: 120, loss: 0.0025786643382161856
step: 130, loss: 0.008905298076570034
step: 140, loss: 0.17424778640270233
step: 150, loss: 0.028896169736981392
step: 160, loss: 0.008175228722393513
step: 170, loss: 0.011499164626002312
step: 180, loss: 0.052407633513212204
step: 190, loss: 0.03668871149420738
step: 200, loss: 0.1859106421470642
step: 210, loss: 0.11486347764730453
/usr/local/lib/python3.7/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=0.7276341948310139, f1=0.6951219512195121,
best_f1=0.7140115163147792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06310772150754929
step: 10, loss: 0.0006117094890214503
step: 20, loss: 0.030899742618203163
step: 30, loss: 0.07412945479154587
step: 40, loss: 0.0325624980032444
step: 50, loss: 0.003974020481109619
step: 60, loss: 0.0009364323923364282
step: 70, loss: 0.02529899775981903

```

```

step: 80, loss: 0.008799763396382332
step: 90, loss: 0.0010940412757918239
step: 100, loss: 0.004245454911142588
step: 110, loss: 0.006957410369068384
step: 120, loss: 0.000645091466140002
step: 130, loss: 0.0009142245398834348
step: 140, loss: 0.035523589700460434
step: 150, loss: 0.0015798226231709123
step: 160, loss: 0.0035226941108703613
step: 170, loss: 0.00645778002217412
step: 180, loss: 0.04540183022618294
step: 190, loss: 0.0006877456326037645
step: 200, loss: 0.0021085271146148443
step: 210, loss: 0.03858080506324768
/usr/local/lib/python3.7/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=0.7256317689530687, f1=0.6901669758812616,
best_f1=0.7140115163147792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011113530490547419
step: 10, loss: 0.02745102532207966
step: 20, loss: 0.004194881767034531
step: 30, loss: 0.00659456942230463
step: 40, loss: 0.09481541067361832
step: 50, loss: 0.018226802349090576
step: 60, loss: 0.0015902523882687092
step: 70, loss: 0.005016448441892862
step: 80, loss: 0.0019209117162972689
step: 90, loss: 0.0204229224473238
step: 100, loss: 0.004031234420835972
step: 110, loss: 0.003709788667038083
step: 120, loss: 0.021580148488283157
step: 130, loss: 0.045992203056812286
step: 140, loss: 0.04945382848381996
step: 150, loss: 0.06813188642263412
step: 160, loss: 0.0023878756910562515
step: 170, loss: 0.007205802947282791
step: 180, loss: 0.03394929692149162
step: 190, loss: 0.013822024688124657
step: 200, loss: 0.0021038600243628025
step: 210, loss: 0.12365009635686874
/usr/local/lib/python3.7/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=0.7266055045871559, f1=0.7024952015355086,

```



```

best_f1=0.7140115163147792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006484907353296876
step: 10, loss: 0.00021352359908632934
step: 20, loss: 0.00418025953695178
step: 30, loss: 0.07185611128807068
step: 40, loss: 0.008801545947790146
step: 50, loss: 0.0007150223245844245
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.002455702982842922
step: 70, loss: 0.003967368043959141
step: 80, loss: 0.006311046425253153
step: 90, loss: 0.02130412496626377
step: 100, loss: 0.00861712172627449
step: 110, loss: 0.0007342399912886322
step: 120, loss: 0.021204523742198944
step: 130, loss: 0.1912127584218979
step: 140, loss: 0.00043280553654767573
step: 150, loss: 0.004838049411773682
step: 160, loss: 0.005590669810771942
step: 170, loss: 0.001411549630574882
step: 180, loss: 0.0009937460999935865
step: 190, loss: 0.017548855394124985
step: 200, loss: 0.013696308247745037
step: 210, loss: 0.013828282244503498
/usr/local/lib/python3.7/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=0.7253141831238779, f1=0.6951672862453533,
best_f1=0.7140115163147792
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:05, 448.99it/s]
load_f1 = 0.7500000000000001
real_f1 = 0.7504621072088725
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, 241.75it/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: 0.5439402461051941
step: 10, loss: 0.36516064405441284
step: 20, loss: 0.2947956323623657
step: 30, loss: 0.40427908301353455
step: 40, loss: 0.42080286145210266
step: 50, loss: 0.30683618783950806
step: 60, loss: 0.2668924331665039
step: 70, loss: 0.26196345686912537
step: 80, loss: 0.26565778255462646
step: 90, loss: 0.26432350277900696
step: 100, loss: 0.374233216047287
step: 110, loss: 0.43629637360572815
step: 120, loss: 0.1393444687128067
step: 130, loss: 0.12166440486907959
step: 140, loss: 0.056418538093566895
step: 150, loss: 0.2001105397939682
step: 160, loss: 0.07902207970619202
step: 170, loss: 0.16831190884113312
step: 180, loss: 0.026209568604826927
step: 190, loss: 0.18568174540996552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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 1: dev_f1=0.7466666666666666, f1=0.7506849315068493,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1939963698387146
step: 10, loss: 0.09317220747470856
step: 20, loss: 0.11393474042415619
step: 30, loss: 0.23118694126605988
step: 40, loss: 0.15486638247966766
step: 50, loss: 0.06577528268098831
step: 60, loss: 0.13862884044647217
step: 70, loss: 0.18440692126750946
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.24585525691509247
step: 90, loss: 0.19554933905601501
step: 100, loss: 0.015587177127599716
step: 110, loss: 0.10795698314905167
step: 120, loss: 0.10014928877353668
step: 130, loss: 0.07039670646190643

```

```

step: 140, loss: 0.054934266954660416
step: 150, loss: 0.09448693692684174
step: 160, loss: 0.02844833768904209
step: 170, loss: 0.1535429060459137
step: 180, loss: 0.05684565380215645
step: 190, loss: 0.025501472875475883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7903614457831326, f1=0.8048780487804879,
best_f1=0.8048780487804879
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032181814312934875
step: 10, loss: 0.03657582029700279
step: 20, loss: 0.04235700145363808
step: 30, loss: 0.0071753705851733685
step: 40, loss: 0.031109651550650597
step: 50, loss: 0.2561097741127014
step: 60, loss: 0.033992502838373184
step: 70, loss: 0.14554613828659058
step: 80, loss: 0.10717020183801651
step: 90, loss: 0.04242190718650818
step: 100, loss: 0.03980671986937523
step: 110, loss: 0.08651067316532135
step: 120, loss: 0.018010737374424934
step: 130, loss: 0.006337661296129227
step: 140, loss: 0.015294566750526428
step: 150, loss: 0.1220286637544632
step: 160, loss: 0.030063867568969727
step: 170, loss: 0.07611624151468277
step: 180, loss: 0.015225681476294994
step: 190, loss: 0.009196407161653042
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.810344827586207, f1=0.7988338192419825,
best_f1=0.7988338192419825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017001943662762642
step: 10, loss: 0.09645450115203857
step: 20, loss: 0.04817387834191322
step: 30, loss: 0.07957509905099869
step: 40, loss: 0.019034013152122498
step: 50, loss: 0.011134213767945766
step: 60, loss: 0.02981785498559475
step: 70, loss: 0.03343730419874191
step: 80, loss: 0.25227561593055725
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.0032673650421202183
step: 100, loss: 0.1651405692100525
step: 110, loss: 0.006630351301282644
step: 120, loss: 0.01578984409570694
step: 130, loss: 0.09061995148658752
step: 140, loss: 0.02113441936671734
step: 150, loss: 0.12256892770528793
step: 160, loss: 0.02868792600929737
step: 170, loss: 0.0611596517264843
step: 180, loss: 0.016731245443224907
step: 190, loss: 0.20376457273960114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8140703517587939, f1=0.8123393316195373,
best_f1=0.8123393316195373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10752535611391068
step: 10, loss: 0.0016181230312213302
step: 20, loss: 0.11714883893728256
step: 30, loss: 0.0022778534330427647
step: 40, loss: 0.05570357292890549
step: 50, loss: 0.052159346640110016
step: 60, loss: 0.039830271154642105
step: 70, loss: 0.07103849947452545
step: 80, loss: 0.06963849812746048
step: 90, loss: 0.03149787709116936
step: 100, loss: 0.002367486944422126
step: 110, loss: 0.05224151536822319
step: 120, loss: 0.006356357131153345

```

```

step: 130, loss: 0.05656808614730835
step: 140, loss: 0.016245199367403984
step: 150, loss: 0.002361846389248967
step: 160, loss: 0.009197820909321308
step: 170, loss: 0.04943336918950081
step: 180, loss: 0.004858407191932201
step: 190, loss: 0.062273118644952774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8150000000000001, f1=0.8232323232323232,
best_f1=0.8232323232323232
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003612353466451168
step: 10, loss: 0.15986336767673492
step: 20, loss: 0.06426452100276947
step: 30, loss: 0.024783533066511154
step: 40, loss: 0.15452276170253754
step: 50, loss: 0.0272079911082983
step: 60, loss: 0.011313538998365402
step: 70, loss: 0.013397290371358395
step: 80, loss: 0.10083053261041641
step: 90, loss: 0.030777955427765846
step: 100, loss: 0.0007355530397035182
step: 110, loss: 0.23622767627239227
step: 120, loss: 0.001839775824919343
step: 130, loss: 0.07763588428497314
step: 140, loss: 0.028709691017866135
step: 150, loss: 0.013212183490395546
step: 160, loss: 0.054643698036670685
step: 170, loss: 0.03797474130988121
step: 180, loss: 0.013114810921251774
step: 190, loss: 0.0066324458457529545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8346883468834688, f1=0.8206521739130436,
best_f1=0.8206521739130436

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001033325563184917
step: 10, loss: 0.12322895228862762
step: 20, loss: 0.058424077928066254
step: 30, loss: 0.01584329828619957
step: 40, loss: 0.0023380930069833994
step: 50, loss: 0.05317375808954239
step: 60, loss: 0.03098316304385662
step: 70, loss: 0.01993187516927719
step: 80, loss: 0.19274918735027313
step: 90, loss: 0.0044153062626719475
step: 100, loss: 0.0005547028267756104
step: 110, loss: 0.012760755605995655
step: 120, loss: 0.0024218950420618057
step: 130, loss: 0.0024549926165491343
step: 140, loss: 0.11940626055002213
step: 150, loss: 0.08843296021223068
step: 160, loss: 0.06724132597446442
step: 170, loss: 0.007661111652851105
step: 180, loss: 0.0018244589446112514
step: 190, loss: 0.009835930541157722
/usr/local/lib/python3.7/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=0.792079207920792, f1=0.7970297029702971,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14749819040298462
step: 10, loss: 0.04398613050580025
step: 20, loss: 0.0017582665896043181
step: 30, loss: 0.001266029430553317
step: 40, loss: 0.010532218962907791
step: 50, loss: 0.010418361984193325
step: 60, loss: 0.4968297481536865
step: 70, loss: 0.004211992956697941
step: 80, loss: 0.000454182387329638
step: 90, loss: 0.0500878281891346
step: 100, loss: 0.08173678070306778
step: 110, loss: 0.0023962499108165503
step: 120, loss: 0.0025736824609339237
step: 130, loss: 0.00087694579269737
step: 140, loss: 0.009013372473418713
step: 150, loss: 0.003255080198869109
step: 160, loss: 0.0033084815368056297

```



```

step: 170, loss: 0.0022272663190960884
step: 180, loss: 0.004251781385391951
step: 190, loss: 0.11870266497135162
/usr/local/lib/python3.7/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=0.8219895287958114, f1=0.8150134048257373,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004626001231372356
step: 10, loss: 0.06685422360897064
step: 20, loss: 0.0009430694626644254
step: 30, loss: 0.0027815629728138447
step: 40, loss: 0.0023971416521817446
step: 50, loss: 0.0012972034746780992
step: 60, loss: 0.0005004821578040719
step: 70, loss: 0.006638979539275169
step: 80, loss: 0.001069654361344874
step: 90, loss: 0.017930204048752785
step: 100, loss: 0.03189118206501007
step: 110, loss: 0.01628296636044979
step: 120, loss: 0.0024353498592972755
step: 130, loss: 0.03815178945660591
step: 140, loss: 0.005738279316574335
step: 150, loss: 0.0005184822948649526
step: 160, loss: 0.027732372283935547
step: 170, loss: 0.0027398578822612762
step: 180, loss: 0.06687425822019577
step: 190, loss: 0.00045872924965806305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8222811671087534, f1=0.8390501319261214,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008002299000509083
step: 10, loss: 0.0026545561850070953
step: 20, loss: 0.01171979308128357
step: 30, loss: 0.0003424487076699734
step: 40, loss: 0.06547130644321442
step: 50, loss: 0.0033441041596233845
step: 60, loss: 0.003236376214772463
step: 70, loss: 0.005363161675632
step: 80, loss: 0.0025948435068130493

```

```

step: 90, loss: 0.0029329194221645594
step: 100, loss: 0.0007120597874745727
step: 110, loss: 0.025411004200577736
step: 120, loss: 0.06200795620679855
step: 130, loss: 0.02440105751156807
step: 140, loss: 0.0005239486927166581
step: 150, loss: 0.00014623133756686002
step: 160, loss: 0.0005873229238204658
step: 170, loss: 0.005873252637684345
step: 180, loss: 0.0003092579136136919
step: 190, loss: 0.002270437078550458
/usr/local/lib/python3.7/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=0.8113695090439276, f1=0.8219895287958114,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000669904809910804
step: 10, loss: 0.004160465206950903
step: 20, loss: 0.021444983780384064
step: 30, loss: 0.027173399925231934
step: 40, loss: 0.000867775350343436
step: 50, loss: 0.07205217331647873
step: 60, loss: 0.0016371732344850898
step: 70, loss: 0.00028627520077861845
step: 80, loss: 0.0004334269615355879
step: 90, loss: 0.00032447781995870173
step: 100, loss: 0.0010583067778497934
step: 110, loss: 0.0011366321705281734
step: 120, loss: 0.020112983882427216
step: 130, loss: 0.0016274515073746443
step: 140, loss: 0.0005382033996284008
step: 150, loss: 0.0008784653036855161
step: 160, loss: 0.0022421330213546753
step: 170, loss: 0.0003436567494645715
step: 180, loss: 0.0003535392170306295
step: 190, loss: 0.0003506252251099795
/usr/local/lib/python3.7/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=0.8205128205128205, f1=0.837696335078534,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016654012142680585

```

```
step: 10, loss: 0.0030431607738137245
step: 20, loss: 0.0003404426679480821
step: 30, loss: 0.03672299534082413
step: 40, loss: 0.010936412960290909
step: 50, loss: 0.006059952080249786
step: 60, loss: 0.014667518436908722
step: 70, loss: 0.0007351799285970628
step: 80, loss: 0.0004388221714179963
step: 90, loss: 0.0002627584617584944
step: 100, loss: 0.00038925063563510776
step: 110, loss: 0.0010809123050421476
step: 120, loss: 0.0001959095970960334
step: 130, loss: 0.0188855342566967
step: 140, loss: 0.20185993611812592
step: 150, loss: 0.00032943731639534235
step: 160, loss: 0.002925880951806903
step: 170, loss: 0.0005322684301063418
step: 180, loss: 0.00026843181694857776
step: 190, loss: 0.0005698181339539587
/usr/local/lib/python3.7/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=0.8195876288659794, f1=0.83289124668435,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008025470888242126
step: 10, loss: 0.001982124289497733
step: 20, loss: 0.0006456774426624179
step: 30, loss: 0.0021641880739480257
step: 40, loss: 0.009211700409650803
step: 50, loss: 0.2145989090204239
step: 60, loss: 0.0026937942020595074
step: 70, loss: 0.002901275409385562
step: 80, loss: 0.021041331812739372
step: 90, loss: 0.04145197570323944
step: 100, loss: 0.03625527396798134
step: 110, loss: 0.0004056654288433492
step: 120, loss: 0.0022752778604626656
step: 130, loss: 0.0005223251646384597
step: 140, loss: 0.00037690013414248824
step: 150, loss: 0.00022016091679688543
step: 160, loss: 0.0004348215297795832
step: 170, loss: 0.006558214779943228
step: 180, loss: 0.005879868287593126
step: 190, loss: 0.026114827021956444
/usr/local/lib/python3.7/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=0.8123393316195373, f1=0.8368421052631578,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014602325391024351
step: 10, loss: 0.0007115734624676406
step: 20, loss: 0.0005387397250160575
step: 30, loss: 0.00033377078943885863
step: 40, loss: 0.001433036057278514
step: 50, loss: 0.0005760076455771923
step: 60, loss: 0.004312198609113693
step: 70, loss: 0.0016067592659965158
step: 80, loss: 0.0007289675413630903
step: 90, loss: 0.003539101220667362
step: 100, loss: 0.0069495197385549545
step: 110, loss: 0.0007645214209333062
step: 120, loss: 0.0007324680336751044
step: 130, loss: 0.05119573324918747
step: 140, loss: 0.0011851510498672724
step: 150, loss: 0.000904404150787741
step: 160, loss: 0.0063232360407710075
step: 170, loss: 0.004009540192782879
step: 180, loss: 0.00829368457198143
step: 190, loss: 0.002698339056223631
/usr/local/lib/python3.7/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=0.8165374677002584, f1=0.8263157894736842,
best_f1=0.8206521739130436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022486839443445206
step: 10, loss: 0.0007133895996958017
step: 20, loss: 0.0005999365821480751
step: 30, loss: 0.0004976121126674116
step: 40, loss: 0.00024195735750254244
step: 50, loss: 0.0016350290970876813
step: 60, loss: 0.0009040149743668735
step: 70, loss: 0.018931785598397255
step: 80, loss: 0.05496843159198761
step: 90, loss: 0.000584329420235008
step: 100, loss: 0.0027944790199398994
step: 110, loss: 0.0006128045497462153
step: 120, loss: 0.000530969409737736

```

```

step: 130, loss: 0.0034981181379407644
step: 140, loss: 0.049600616097450256
step: 150, loss: 0.0009030324872583151
step: 160, loss: 0.00027549214428290725
step: 170, loss: 0.00040512304985895753
step: 180, loss: 0.00037986168172210455
step: 190, loss: 0.0005730355042032897
/usr/local/lib/python3.7/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=0.8152173913043478, f1=0.8356545961002786,
best_f1=0.8206521739130436
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, 224.53it/s]
load_f1 = 0.7405541561712846
real_f1 = 0.7448979591836734
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:18, 238.48it/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: 0.6422535181045532
step: 10, loss: 0.3855610191822052
step: 20, loss: 0.3080427050590515
step: 30, loss: 0.3724243640899658
step: 40, loss: 0.28648874163627625
step: 50, loss: 0.2804260849952698
step: 60, loss: 0.2788405418395996
step: 70, loss: 0.3618849217891693
step: 80, loss: 0.33976060152053833
step: 90, loss: 0.23184998333454132
step: 100, loss: 0.36698439717292786
step: 110, loss: 0.19723403453826904
step: 120, loss: 0.2112935185432434
step: 130, loss: 0.01680266112089157
step: 140, loss: 0.28490301966667175
step: 150, loss: 0.277375191450119
step: 160, loss: 0.12757699191570282
step: 170, loss: 0.18078164756298065

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7259615384615385, f1=0.711943793911007,
best_f1=0.711943793911007
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04015108942985535
step: 10, loss: 0.08797676861286163
step: 20, loss: 0.059039074927568436
step: 30, loss: 0.20474332571029663
step: 40, loss: 0.09339619427919388
step: 50, loss: 0.13118501007556915
step: 60, loss: 0.07540220767259598
step: 70, loss: 0.10836632549762726
step: 80, loss: 0.029277004301548004
step: 90, loss: 0.08953448385000229
step: 100, loss: 0.2230033427476883
step: 110, loss: 0.0896873027086258
step: 120, loss: 0.06025608628988266
step: 130, loss: 0.11858248710632324
step: 140, loss: 0.3143608868122101
step: 150, loss: 0.1326310932636261
step: 160, loss: 0.24243733286857605
step: 170, loss: 0.197067528963089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8100686498855835, 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: 0.07398577779531479
step: 10, loss: 0.0792209580540657
step: 20, loss: 0.012533504515886307
step: 30, loss: 0.05502643436193466
step: 40, loss: 0.1734725385904312
step: 50, loss: 0.18224777281284332

```



```

step: 60, loss: 0.04492606595158577
step: 70, loss: 0.0733509436249733
step: 80, loss: 0.041022732853889465
step: 90, loss: 0.11008584499359131
step: 100, loss: 0.1278122216463089
step: 110, loss: 0.1057194247841835
step: 120, loss: 0.042282044887542725
step: 130, loss: 0.04311869293451309
step: 140, loss: 0.0507475920021534
step: 150, loss: 0.011861332692205906
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.03574354201555252
step: 170, loss: 0.0424148328602314
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8321513002364066, f1=0.8129330254041571,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008846912533044815
step: 10, loss: 0.024498043581843376
step: 20, loss: 0.03279440477490425
step: 30, loss: 0.048995427787303925
step: 40, loss: 0.0015687108971178532
step: 50, loss: 0.15592046082019806
step: 60, loss: 0.14982165396213531
step: 70, loss: 0.00525631895288825
step: 80, loss: 0.05856805294752121
step: 90, loss: 0.04216037318110466
step: 100, loss: 0.09444602578878403
step: 110, loss: 0.05436166375875473
step: 120, loss: 0.0034433810506016016
step: 130, loss: 0.0399927943944931
step: 140, loss: 0.040335483849048615
step: 150, loss: 0.1775304228067398
step: 160, loss: 0.08909899741411209
step: 170, loss: 0.0012113817501813173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8377723970944311, f1=0.8124999999999999,
best_f1=0.8124999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014803150668740273
step: 10, loss: 0.0067827822640538216
step: 20, loss: 0.008187858387827873
step: 30, loss: 0.0919446274638176
step: 40, loss: 0.0033208990935236216
step: 50, loss: 0.13695724308490753
step: 60, loss: 0.01162189431488514
step: 70, loss: 0.23315756022930145
step: 80, loss: 0.034918420016765594
step: 90, loss: 0.1460004597902298
step: 100, loss: 0.021642332896590233
step: 110, loss: 0.038768403232097626
step: 120, loss: 0.034246329218149185
step: 130, loss: 0.043001938611269
step: 140, loss: 0.02607409469783306
step: 150, loss: 0.09063055366277695
step: 160, loss: 0.04265226051211357
step: 170, loss: 0.006479672156274319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8423529411764706, f1=0.817351598173516,
best_f1=0.817351598173516
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002296297810971737
step: 10, loss: 0.017569804564118385
step: 20, loss: 0.008984273299574852
step: 30, loss: 0.01778661459684372
step: 40, loss: 0.047750845551490784
step: 50, loss: 0.1543836146593094
step: 60, loss: 0.03084374964237213
step: 70, loss: 0.10486926883459091
step: 80, loss: 0.009651917032897472
step: 90, loss: 0.06207887828350067
step: 100, loss: 0.017436433583498

```

```

step: 110, loss: 0.013464096002280712
step: 120, loss: 0.0037958966568112373
step: 130, loss: 0.033238235861063004
step: 140, loss: 0.004450452979654074
step: 150, loss: 0.030815180391073227
step: 160, loss: 0.21448209881782532
step: 170, loss: 0.013848892413079739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8551068883610452, f1=0.8309859154929576,
best_f1=0.8309859154929576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006999194156378508
step: 10, loss: 0.001281821052543819
step: 20, loss: 0.005479228682816029
step: 30, loss: 0.01370985247194767
step: 40, loss: 0.09229680150747299
step: 50, loss: 0.001772531308233738
step: 60, loss: 0.0054123359732329845
step: 70, loss: 0.0006071723182685673
step: 80, loss: 0.0273466557264328
step: 90, loss: 0.0002897186204791069
step: 100, loss: 0.00024694009334780276
step: 110, loss: 0.009194407612085342
step: 120, loss: 0.015940720215439796
step: 130, loss: 0.1358327865600586
step: 140, loss: 0.002368451561778784
step: 150, loss: 0.009239648468792439
step: 160, loss: 0.0025312439538538456
step: 170, loss: 0.003690927755087614
/usr/local/lib/python3.7/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=0.8421052631578947, f1=0.8280871670702179,
best_f1=0.8309859154929576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018267922103405
step: 10, loss: 0.011733087711036205
step: 20, loss: 0.000760805734898895

```

```

step: 30, loss: 0.025286275893449783
step: 40, loss: 0.0005136149120517075
step: 50, loss: 0.0017104905564337969
step: 60, loss: 0.0028559656348079443
step: 70, loss: 0.07535115629434586
step: 80, loss: 0.001465153181925416
step: 90, loss: 0.002059033140540123
step: 100, loss: 0.042088426649570465
step: 110, loss: 0.08443055301904678
step: 120, loss: 0.0017574153607711196
step: 130, loss: 0.0042371745221316814
step: 140, loss: 0.08994156122207642
step: 150, loss: 0.04844025894999504
step: 160, loss: 0.0224827341735363
step: 170, loss: 0.06593939661979675
/usr/local/lib/python3.7/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=0.8153846153846153, f1=0.8402948402948404,
best_f1=0.8309859154929576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06422460824251175
step: 10, loss: 0.006243431940674782
step: 20, loss: 0.10536587983369827
step: 30, loss: 0.0026446289848536253
step: 40, loss: 0.017057381570339203
step: 50, loss: 0.00492074154317379
step: 60, loss: 0.005643489770591259
step: 70, loss: 0.006952753756195307
step: 80, loss: 0.000683940015733242
step: 90, loss: 0.05618295818567276
step: 100, loss: 0.07418564707040787
step: 110, loss: 0.05306246876716614
step: 120, loss: 0.00799037329852581
step: 130, loss: 0.06521525979042053
step: 140, loss: 0.058148518204689026
step: 150, loss: 0.007405312731862068
step: 160, loss: 0.038275495171546936
step: 170, loss: 0.004579536616802216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8557692307692307, f1=0.8329411764705882,
best_f1=0.8329411764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03410353884100914
step: 10, loss: 0.00023967644665390253
step: 20, loss: 0.05234919488430023
step: 30, loss: 0.002587870927527547
step: 40, loss: 0.00011729048856068403
step: 50, loss: 0.14194051921367645
step: 60, loss: 0.00011926968727493659
step: 70, loss: 0.011202145367860794
step: 80, loss: 0.0022552923765033484
step: 90, loss: 0.006901460699737072
step: 100, loss: 9.836380195338279e-05
step: 110, loss: 0.001807809341698885
step: 120, loss: 0.0010378692531958222
step: 130, loss: 0.008074501529335976
step: 140, loss: 0.0325695239007473
step: 150, loss: 0.032715097069740295
step: 160, loss: 0.0039015556685626507
step: 170, loss: 0.00043241865932941437
/usr/local/lib/python3.7/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=0.8511749347258485, f1=0.8463476070528967,
best_f1=0.8329411764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05032636597752571
step: 10, loss: 0.007221690379083157
step: 20, loss: 0.0004037350881844759
step: 30, loss: 0.0002195376728195697
step: 40, loss: 0.04820990934967995
step: 50, loss: 0.0015187592944130301
step: 60, loss: 0.014651377685368061
step: 70, loss: 0.00039710450801067054
step: 80, loss: 0.00029527719016186893
step: 90, loss: 0.013670805841684341
step: 100, loss: 0.010514293797314167
step: 110, loss: 0.005694935563951731
step: 120, loss: 0.0011588007910177112
step: 130, loss: 0.00018876649846788496
step: 140, loss: 0.0006930554518476129
step: 150, loss: 0.008536371402442455
step: 160, loss: 0.007665439508855343

```

```

step: 170, loss: 0.003972303122282028
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.87, f1=0.8292682926829268, best_f1=0.8292682926829268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012137656100094318
step: 10, loss: 0.00031039875466376543
step: 20, loss: 0.0001266394683625549
step: 30, loss: 0.0006267053540796041
step: 40, loss: 0.0006407922483049333
step: 50, loss: 0.00010979323997162282
step: 60, loss: 0.004250267054885626
step: 70, loss: 0.08450960367918015
step: 80, loss: 8.457822696072981e-05
step: 90, loss: 0.0003225612745154649
step: 100, loss: 0.000992592773400247
step: 110, loss: 0.006547757890075445
step: 120, loss: 0.15318721532821655
step: 130, loss: 0.008009916171431541
step: 140, loss: 0.01749350130558014
step: 150, loss: 0.03808460384607315
step: 160, loss: 0.0010417295852676034
step: 170, loss: 0.00040537663153372705
/usr/local/lib/python3.7/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=0.8564705882352941, f1=0.8158508158508158,
best_f1=0.8292682926829268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025727376341819763
step: 10, loss: 0.0012553018750622869
step: 20, loss: 0.00272872531786561
step: 30, loss: 0.0005363799864426255
step: 40, loss: 0.007675267290323973
step: 50, loss: 0.0006180663476698101
step: 60, loss: 0.020625442266464233
step: 70, loss: 0.0015291298041120172
step: 80, loss: 0.009401723742485046
step: 90, loss: 9.008869528770447e-05

```

```

step: 100, loss: 0.001406209310516715
step: 110, loss: 0.031987786293029785
step: 120, loss: 0.030583010986447334
step: 130, loss: 0.0006700814701616764
step: 140, loss: 0.0044068326242268085
step: 150, loss: 0.1486494094133377
step: 160, loss: 0.0018793969647958875
step: 170, loss: 0.009997940622270107
/usr/local/lib/python3.7/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=0.8528678304239402, f1=0.8557213930348259,
best_f1=0.8292682926829268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007805326953530312
step: 10, loss: 0.00011873827315866947
step: 20, loss: 0.0002016277430811897
step: 30, loss: 0.0007619487587362528
step: 40, loss: 0.00019500296912156045
step: 50, loss: 0.0008085292065516114
step: 60, loss: 0.0005472433404065669
step: 70, loss: 0.0009943434270098805
step: 80, loss: 0.002503411378711462
step: 90, loss: 0.0003248518332839012
step: 100, loss: 0.003888699458912015
step: 110, loss: 0.029873333871364594
step: 120, loss: 0.042630087584257126
step: 130, loss: 0.012387774884700775
step: 140, loss: 0.013734773732721806
step: 150, loss: 0.0006796930683776736
step: 160, loss: 0.00020543420396279544
step: 170, loss: 0.0008524444419890642
/usr/local/lib/python3.7/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=0.8629441624365481, f1=0.85785536159601,
best_f1=0.8292682926829268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012002446601400152
step: 10, loss: 0.005381309427320957
step: 20, loss: 0.017831770703196526
step: 30, loss: 0.0003658268542494625
step: 40, loss: 0.0026316503062844276
step: 50, loss: 0.0008382871747016907

```

```

step: 60, loss: 0.005512910895049572
step: 70, loss: 0.000574883131776005
step: 80, loss: 0.02583920769393444
step: 90, loss: 0.00038059233338572085
step: 100, loss: 0.001116046216338873
step: 110, loss: 0.0002722761419136077
step: 120, loss: 0.00047265170724131167
step: 130, loss: 0.003444971051067114
step: 140, loss: 0.013440349139273167
step: 150, loss: 0.0003178475017193705
step: 160, loss: 0.0008839200017973781
step: 170, loss: 0.011856055818498135
/usr/local/lib/python3.7/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=0.8656716417910448, f1=0.8410757946210269,
best_f1=0.8292682926829268
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, 265.49it/s]

```



```
load_f1 = 0.4009111617312073
real_f1 = 0.3919372900335946
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, 237.91it/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 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: 0.6034992933273315
step: 10, loss: 0.6254540085792542
step: 20, loss: 0.3161802589893341
step: 30, loss: 0.11813627183437347
step: 40, loss: 0.23216187953948975
step: 50, loss: 0.10539103299379349
step: 60, loss: 0.034173015505075455
step: 70, loss: 0.04667958989739418
step: 80, loss: 0.07395090907812119
step: 90, loss: 0.13011881709098816
step: 100, loss: 0.004187609069049358
step: 110, loss: 0.16106760501861572
step: 120, loss: 0.010089356452226639
```

```

step: 130, loss: 0.015396957285702229
step: 140, loss: 0.0025474426802247763
step: 150, loss: 0.010820802301168442
step: 160, loss: 0.014039781875908375
step: 170, loss: 0.17835940420627594
step: 180, loss: 0.011064998805522919
step: 190, loss: 0.05278492718935013
step: 200, loss: 0.08992224931716919
step: 210, loss: 0.005397498607635498
step: 220, loss: 0.07936793565750122
step: 230, loss: 0.012837054207921028
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9776286353467561, f1=0.9684684684684683,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008343212306499481
step: 10, loss: 0.008372928015887737
step: 20, loss: 0.10709783434867859
step: 30, loss: 0.3013731837272644
step: 40, loss: 0.034365326166152954
step: 50, loss: 0.0025941089261323214
step: 60, loss: 0.022047143429517746
step: 70, loss: 0.10936609655618668
step: 80, loss: 0.03582875803112984
step: 90, loss: 0.07913580536842346
step: 100, loss: 0.04599151015281677
step: 110, loss: 0.0252387672662735
step: 120, loss: 0.00445564417168498
step: 130, loss: 0.004343729466199875
step: 140, loss: 0.00422599958255887
step: 150, loss: 0.045258715748786926
step: 160, loss: 0.06699656695127487
step: 170, loss: 0.0016609459416940808
step: 180, loss: 0.00413130410015583
step: 190, loss: 0.015527930110692978
step: 200, loss: 0.00151492387522012
step: 210, loss: 0.0009388602920807898
step: 220, loss: 0.16525451838970184
step: 230, loss: 0.003136218059808016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9799107142857142, f1=0.9707865168539327,
best_f1=0.9707865168539327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01769454963505268
step: 10, loss: 0.01878429390490055
step: 20, loss: 0.0060164486058056355
step: 30, loss: 0.026026563718914986
step: 40, loss: 0.11805757880210876
step: 50, loss: 0.0037975965533405542
step: 60, loss: 0.00367822521366179
step: 70, loss: 0.026092294603586197
step: 80, loss: 0.0016156204510480165
step: 90, loss: 0.028417635709047318
step: 100, loss: 0.0013770076911896467
step: 110, loss: 0.0007329051150009036
step: 120, loss: 0.02779513970017433
step: 130, loss: 0.0007971915183588862
step: 140, loss: 0.008639758452773094
step: 150, loss: 0.018455177545547485
step: 160, loss: 0.05510491877794266
step: 170, loss: 0.040010493248701096
step: 180, loss: 0.012836846522986889
step: 190, loss: 0.0006512771942652762
step: 200, loss: 0.0010500815697014332
step: 210, loss: 0.034754686057567596
step: 220, loss: 0.0004944994580000639
step: 230, loss: 0.03360074386000633
/usr/local/lib/python3.7/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=0.9778761061946903, f1=0.9700332963374029,
best_f1=0.9707865168539327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006558027234859765
step: 10, loss: 0.0018433601362630725
step: 20, loss: 0.003262127749621868
step: 30, loss: 0.0004989276640117168
step: 40, loss: 0.003957807086408138

```

```

step: 50, loss: 0.001463598571717739
step: 60, loss: 0.000990512315183878
step: 70, loss: 0.03842731937766075
step: 80, loss: 0.007152171805500984
step: 90, loss: 0.004477944225072861
step: 100, loss: 0.0009279890218749642
step: 110, loss: 0.0010740832658484578
step: 120, loss: 0.11167839169502258
step: 130, loss: 0.0018084882758557796
step: 140, loss: 0.002984248101711273
step: 150, loss: 0.12446165084838867
step: 160, loss: 0.13727273046970367
step: 170, loss: 0.004629857372492552
step: 180, loss: 0.0008791365544311702
step: 190, loss: 0.0014497919473797083
step: 200, loss: 0.0014770784182474017
step: 210, loss: 0.13937076926231384
step: 220, loss: 0.0006397421821020544
step: 230, loss: 0.005889765918254852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831271091113611, f1=0.9762174405436014,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008191895321942866
step: 10, loss: 0.0007335151894949377
step: 20, loss: 0.0008552615763619542
step: 30, loss: 0.00037427834467962384
step: 40, loss: 0.0009257504134438932
step: 50, loss: 0.0003572565328795463
step: 60, loss: 0.12794430553913116
step: 70, loss: 0.0009287144057452679
step: 80, loss: 0.0028632287867367268
step: 90, loss: 0.009454711340367794
step: 100, loss: 0.00024858684628270566
step: 110, loss: 0.004539933521300554
step: 120, loss: 0.00014666060451418161
step: 130, loss: 0.0016091479919850826
step: 140, loss: 0.003355347318574786
step: 150, loss: 0.0008816327899694443
step: 160, loss: 0.03935618698596954

```

```

step: 170, loss: 0.009042437188327312
step: 180, loss: 0.001903068390674889
step: 190, loss: 0.11226052790880203
step: 200, loss: 0.027041297405958176
step: 210, loss: 0.00137425702996552
step: 220, loss: 0.0008398644276894629
step: 230, loss: 0.0008547224570065737
/usr/local/lib/python3.7/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=0.9821826280623607, f1=0.9755011135857461,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00655668880790472
step: 10, loss: 0.001225435989908874
step: 20, loss: 0.001194032491184771
step: 30, loss: 0.010243331082165241
step: 40, loss: 0.0001341982715530321
step: 50, loss: 0.0013509695418179035
step: 60, loss: 0.00012278092617634684
step: 70, loss: 0.00014066848962102085
step: 80, loss: 0.0063690319657325745
step: 90, loss: 0.00022583288955502212
step: 100, loss: 0.03778747469186783
step: 110, loss: 0.0033049369230866432
step: 120, loss: 0.007463118527084589
step: 130, loss: 0.00016089410928543657
step: 140, loss: 0.00020732720440719277
step: 150, loss: 0.03610936552286148
step: 160, loss: 0.0007344717741943896
step: 170, loss: 0.0018288076389580965
step: 180, loss: 0.0007863566279411316
step: 190, loss: 0.001379933673888445
step: 200, loss: 0.0003405580355320126
step: 210, loss: 0.0005966299795545638
step: 220, loss: 0.0001208899702760391
step: 230, loss: 0.04875428229570389
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9765363128491621, f1=0.9754464285714286,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008240891620516777

```

```

step: 10, loss: 0.0002184815821237862
step: 20, loss: 0.0006231982260942459
step: 30, loss: 0.00017412535089533776
step: 40, loss: 0.0029284132178872824
step: 50, loss: 0.000800044450443238
step: 60, loss: 0.002524957526475191
step: 70, loss: 0.012071450240910053
step: 80, loss: 0.0026771004777401686
step: 90, loss: 7.7769611380063e-05
step: 100, loss: 0.0001271920627914369
step: 110, loss: 0.0001736471167532727
step: 120, loss: 7.801008905516937e-05
step: 130, loss: 7.876271411078051e-05
step: 140, loss: 0.0003622629155870527
step: 150, loss: 0.007184998132288456
step: 160, loss: 0.09371455758810043
step: 170, loss: 0.0008442070102319121
step: 180, loss: 0.0013112262822687626
step: 190, loss: 0.0007680425769649446
step: 200, loss: 0.03574890270829201
step: 210, loss: 0.00013213424244895577
step: 220, loss: 0.005225168541073799
step: 230, loss: 0.002629936672747135
/usr/local/lib/python3.7/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=0.9821428571428571, f1=0.9755011135857461,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017397008195985109
step: 10, loss: 0.001814147923141718
step: 20, loss: 0.0033134380355477333
step: 30, loss: 0.0006609224365092814
step: 40, loss: 0.007495000958442688
step: 50, loss: 0.001070731203071773
step: 60, loss: 0.009093350730836391
step: 70, loss: 0.00016061561473179609
step: 80, loss: 0.0005724338116124272
step: 90, loss: 0.00037215143674984574
step: 100, loss: 0.0003908348153345287
step: 110, loss: 0.12198679894208908
step: 120, loss: 0.03094794787466526
step: 130, loss: 0.004334759898483753
step: 140, loss: 0.00012054570106556639
step: 150, loss: 0.0005281816702336073
step: 160, loss: 0.0005078893736936152

```

```

step: 170, loss: 0.0059981802478432655
step: 180, loss: 0.0009744130074977875
step: 190, loss: 0.0010079792700707912
step: 200, loss: 0.04804025590419769
step: 210, loss: 0.002417234005406499
step: 220, loss: 0.0027655374724417925
step: 230, loss: 0.0005453393096104264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854423292273236, f1=0.9732739420935412,
best_f1=0.9732739420935412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018658931367099285
step: 10, loss: 0.0008417811477556825
step: 20, loss: 0.0006596613093279302
step: 30, loss: 8.520358824171126e-05
step: 40, loss: 0.040191952139139175
step: 50, loss: 7.211968477349728e-05
step: 60, loss: 0.0004051201103720814
step: 70, loss: 6.799584662076086e-05
step: 80, loss: 0.00011808897397713736
step: 90, loss: 0.00030717888148501515
step: 100, loss: 0.0005397984641604125
step: 110, loss: 6.367493915604427e-05
step: 120, loss: 7.268620538525283e-05
step: 130, loss: 9.933332330547273e-05
step: 140, loss: 5.781168874818832e-05
step: 150, loss: 0.0014924465212970972
step: 160, loss: 6.362973363138735e-05
step: 170, loss: 0.00028098010807298124
step: 180, loss: 0.0003704625414684415
step: 190, loss: 0.0014313021674752235
step: 200, loss: 4.39264222222846e-05
step: 210, loss: 0.005632465239614248
step: 220, loss: 3.987029049312696e-05
step: 230, loss: 7.02341494616121e-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=0.9853768278965129, f1=0.9763779527559054,
best_f1=0.9732739420935412

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.004643026040867e-05
step: 10, loss: 4.3274321797071025e-05
step: 20, loss: 0.00278400513343513
step: 30, loss: 0.00022984000679571182
step: 40, loss: 7.315370748983696e-05
step: 50, loss: 0.00018247142725158483
step: 60, loss: 0.001954844454303384
step: 70, loss: 8.771685679676011e-05
step: 80, loss: 0.0031757664401084185
step: 90, loss: 0.005097543820738792
step: 100, loss: 8.507190796080977e-05
step: 110, loss: 0.0003270567976869643
step: 120, loss: 0.001557918731123209
step: 130, loss: 0.000123590201837942
step: 140, loss: 0.020959431305527687
step: 150, loss: 0.019468512386083603
step: 160, loss: 3.5667591873789206e-05
step: 170, loss: 4.131372406845912e-05
step: 180, loss: 7.920922507764772e-05
step: 190, loss: 0.013848268426954746
step: 200, loss: 3.1459076126338914e-05
step: 210, loss: 3.268046202720143e-05
step: 220, loss: 3.7996444007148966e-05
step: 230, loss: 0.00019149719446431845
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9765886287625419,
best_f1=0.9732739420935412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000455571775091812
step: 10, loss: 5.623268225463107e-05
step: 20, loss: 4.2555981053737924e-05
step: 30, loss: 0.0013969402061775327
step: 40, loss: 7.290779467439279e-05
step: 50, loss: 0.00017345482774544507
step: 60, loss: 0.0007071066065691411
step: 70, loss: 0.0004418269672896713
step: 80, loss: 4.693949449574575e-05
step: 90, loss: 5.386635893955827e-05
step: 100, loss: 6.381152343237773e-05
step: 110, loss: 0.0054458496160805225
step: 120, loss: 0.00011960989650106058

```

```

step: 130, loss: 4.947110573993996e-05
step: 140, loss: 0.0002637383877299726
step: 150, loss: 0.04006156325340271
step: 160, loss: 0.00010183455742662773
step: 170, loss: 0.0183399748057127
step: 180, loss: 0.00031515603768639266
step: 190, loss: 0.001401746179908514
step: 200, loss: 0.004141676239669323
step: 210, loss: 4.092806193511933e-05
step: 220, loss: 6.042092718416825e-05
step: 230, loss: 0.00013067448162473738
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9764837625979844,
best_f1=0.9732739420935412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.246849367627874e-05
step: 10, loss: 8.21627545519732e-05
step: 20, loss: 0.0001639233814785257
step: 30, loss: 0.00014244441990740597
step: 40, loss: 0.00027628877433016896
step: 50, loss: 0.0023227541241794825
step: 60, loss: 5.579032585956156e-05
step: 70, loss: 0.00026184943271800876
step: 80, loss: 0.00018892042862717062
step: 90, loss: 5.034519926994108e-05
step: 100, loss: 3.06767578877043e-05
step: 110, loss: 3.514296258799732e-05
step: 120, loss: 5.4401170928031206e-05
step: 130, loss: 0.00010186483268626034
step: 140, loss: 0.0016237370437011123
step: 150, loss: 0.0001365121133858338
step: 160, loss: 3.894229666911997e-05
step: 170, loss: 4.914421515422873e-05
step: 180, loss: 0.013828698545694351
step: 190, loss: 0.0008871721802279353
step: 200, loss: 3.201044455636293e-05
step: 210, loss: 7.891379937063903e-05
step: 220, loss: 5.414687257143669e-05
step: 230, loss: 0.03928016871213913
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9787709497206705,
best_f1=0.9732739420935412

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.374080450972542e-05
step: 10, loss: 0.00015122392505872995
step: 20, loss: 7.640974217792973e-05
step: 30, loss: 5.199397855903953e-05
step: 40, loss: 4.155944407102652e-05
step: 50, loss: 0.0005757635226473212
step: 60, loss: 0.0002060357219306752
step: 70, loss: 0.0006681244703941047
step: 80, loss: 6.259779911488295e-05
step: 90, loss: 0.0001872021093731746
step: 100, loss: 5.1749873819062486e-05
step: 110, loss: 0.057711776345968246
step: 120, loss: 0.02027246356010437
step: 130, loss: 0.000358163466444239
step: 140, loss: 3.1187279091682285e-05
step: 150, loss: 4.987998545402661e-05
step: 160, loss: 0.0003367629542481154
step: 170, loss: 5.242121915216558e-05
step: 180, loss: 3.6934568925062194e-05
step: 190, loss: 4.792432810063474e-05
step: 200, loss: 5.45191032870207e-05
step: 210, loss: 1.9542620066204108e-05
step: 220, loss: 2.668350498424843e-05
step: 230, loss: 0.00010458649921929464
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9765886287625419,
best_f1=0.9732739420935412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010038808250101283
step: 10, loss: 7.232600910356268e-05
step: 20, loss: 0.00027419166872277856
step: 30, loss: 0.00013967980339657515
step: 40, loss: 0.010822171345353127
step: 50, loss: 2.8862888939329423e-05
step: 60, loss: 0.00011918189557036385
step: 70, loss: 6.57386626699008e-05
step: 80, loss: 4.46217818534933e-05
step: 90, loss: 7.973342871991917e-05
step: 100, loss: 2.666126238182187e-05
step: 110, loss: 0.0002626903878990561
step: 120, loss: 1.994107515201904e-05

```

```

step: 130, loss: 2.4489587303833105e-05
step: 140, loss: 2.9939326850580983e-05
step: 150, loss: 4.786821955349296e-05
step: 160, loss: 0.010857262648642063
step: 170, loss: 4.142085163039155e-05
step: 180, loss: 0.00019160426745656878
step: 190, loss: 2.739887531788554e-05
step: 200, loss: 2.4414957806584425e-05
step: 210, loss: 0.0015629625413566828
step: 220, loss: 3.0363104087882675e-05
step: 230, loss: 0.03517651930451393
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9776785714285714,
best_f1=0.9732739420935412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.231039252365008e-05
step: 10, loss: 1.716590668365825e-05
step: 20, loss: 2.5905070287990384e-05
step: 30, loss: 8.915459329728037e-05
step: 40, loss: 0.00010997071512974799
step: 50, loss: 0.029951469972729683
step: 60, loss: 3.61671372957062e-05
step: 70, loss: 7.910319254733622e-05
step: 80, loss: 3.430474680499174e-05
step: 90, loss: 0.002203746233135462
step: 100, loss: 4.158885349170305e-05
step: 110, loss: 2.8802925953641534e-05
step: 120, loss: 0.0963287353515625
step: 130, loss: 0.003860436612740159
step: 140, loss: 2.6027983039966784e-05
step: 150, loss: 0.0001724366593407467
step: 160, loss: 2.1289652067935094e-05
step: 170, loss: 2.1907986592850648e-05
step: 180, loss: 0.00014057003136258572
step: 190, loss: 7.412976265186444e-05
step: 200, loss: 3.576516974135302e-05
step: 210, loss: 7.257021934492514e-05
step: 220, loss: 0.00044391743722371757
step: 230, loss: 2.4299530196003616e-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=0.9842696629213483, f1=0.9776785714285714,
best_f1=0.9732739420935412

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 194.67it/s]
load_f1 = 0.9854096520763187
real_f1 = 0.9832026875699889
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, 233.53it/s]
```

4.3.2 DBLP-GoogleScholar - Running the matcher

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

```

--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

Selected optimization level 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: 0.616640567779541
step: 10, loss: 0.5202642679214478
step: 20, loss: 0.46363040804862976
step: 30, loss: 0.10239613801240921
step: 40, loss: 0.17109592258930206
step: 50, loss: 0.12899847328662872
step: 60, loss: 0.08784924447536469
step: 70, loss: 0.05482608079910278
step: 80, loss: 0.03314884006977081
step: 90, loss: 0.1382315307855606
step: 100, loss: 0.03983887657523155
step: 110, loss: 0.06839679926633835
step: 120, loss: 0.08220195025205612
step: 130, loss: 0.15716037154197693
step: 140, loss: 0.1362500637769699
step: 150, loss: 0.032656148076057434
step: 160, loss: 0.0417444184422493
step: 170, loss: 0.1623203605413437
step: 180, loss: 0.034894589334726334
step: 190, loss: 0.007741562556475401
step: 200, loss: 0.14478500187397003
step: 210, loss: 0.0742100179195404
step: 220, loss: 0.18812711536884308
step: 230, loss: 0.13280527293682098
step: 240, loss: 0.06970810890197754
step: 250, loss: 0.020898643881082535
step: 260, loss: 0.15609094500541687
step: 270, loss: 0.039975401014089584
step: 280, loss: 0.04086095467209816
step: 290, loss: 0.030685732141137123
step: 300, loss: 0.029858987778425217

```

```

step: 310, loss: 0.12563425302505493
step: 320, loss: 0.10869754105806351
step: 330, loss: 0.012395375408232212
step: 340, loss: 0.024172451347112656
step: 350, loss: 0.01956765167415142
step: 360, loss: 0.060973893851041794
step: 370, loss: 0.09709060192108154
step: 380, loss: 0.011385143734514713
step: 390, loss: 0.0459306538105011
step: 400, loss: 0.21278177201747894
step: 410, loss: 0.06357576698064804
step: 420, loss: 0.020755160599946976
step: 430, loss: 0.19583216309547424
step: 440, loss: 0.025277521461248398
step: 450, loss: 0.005751447286456823
step: 460, loss: 0.10395904630422592
step: 470, loss: 0.11001613736152649
step: 480, loss: 0.043253690004348755
step: 490, loss: 0.0839887484908104
step: 500, loss: 0.06984424591064453
step: 510, loss: 0.03905481845140457
step: 520, loss: 0.033952195197343826
step: 530, loss: 0.0013559238286688924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9455719557195572, f1=0.9406896551724139,
best_f1=0.9406896551724139
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12083001434803009
step: 10, loss: 0.045880381017923355
step: 20, loss: 0.032555900514125824
step: 30, loss: 0.05153051018714905
step: 40, loss: 0.13679906725883484
step: 50, loss: 0.10045453161001205
step: 60, loss: 0.006198695860803127
step: 70, loss: 0.018849948421120644
step: 80, loss: 0.0718580037355423
step: 90, loss: 0.009122283197939396
step: 100, loss: 0.022609055042266846
step: 110, loss: 0.014955544844269753
step: 120, loss: 0.06588808447122574

```



```

step: 130, loss: 0.17689785361289978
step: 140, loss: 0.047612398862838745
step: 150, loss: 0.08811379224061966
step: 160, loss: 0.009053649380803108
step: 170, loss: 0.009501243941485882
step: 180, loss: 0.018990734592080116
step: 190, loss: 0.06913869082927704
step: 200, loss: 0.008944366127252579
step: 210, loss: 0.039138346910476685
step: 220, loss: 0.0591648668050766
step: 230, loss: 0.021922670304775238
step: 240, loss: 0.022834718227386475
step: 250, loss: 0.07430235296487808
step: 260, loss: 0.001553086913190782
step: 270, loss: 0.01986124739050865
step: 280, loss: 0.009765439666807652
step: 290, loss: 0.011874772608280182
step: 300, loss: 0.16701602935791016
step: 310, loss: 0.013672123663127422
step: 320, loss: 0.06113070994615555
step: 330, loss: 0.030139625072479248
step: 340, loss: 0.013536578044295311
step: 350, loss: 0.0007057343609631062
step: 360, loss: 0.1920817345380783
step: 370, loss: 0.0789003074169159
step: 380, loss: 0.04846622794866562
step: 390, loss: 0.024428214877843857
step: 400, loss: 0.06851480156183243
step: 410, loss: 0.002966918982565403
step: 420, loss: 0.17698709666728973
step: 430, loss: 0.010342120192945004
step: 440, loss: 0.15678183734416962
step: 450, loss: 0.01008677389472723
step: 460, loss: 0.08634123206138611
step: 470, loss: 0.06846003234386444
step: 480, loss: 0.24311380088329315
step: 490, loss: 0.008342727087438107
step: 500, loss: 0.15600422024726868
step: 510, loss: 0.00554746575653553
step: 520, loss: 0.06172148138284683
step: 530, loss: 0.004628555383533239
/usr/local/lib/python3.7/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=0.943155979990905, f1=0.9378120744439402,
best_f1=0.9406896551724139
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03781453147530556
step: 10, loss: 0.028275489807128906
step: 20, loss: 0.056627798825502396
step: 30, loss: 0.11373759061098099
step: 40, loss: 0.032258860766887665
step: 50, loss: 0.05187708139419556
step: 60, loss: 0.014934394508600235
step: 70, loss: 0.0034253871999680996
step: 80, loss: 0.00628720922395587
step: 90, loss: 0.003730400465428829
step: 100, loss: 0.17293433845043182
step: 110, loss: 0.005847099702805281
step: 120, loss: 0.021187355741858482
step: 130, loss: 0.002827863674610853
step: 140, loss: 0.058775417506694794
step: 150, loss: 0.0572192519903183
step: 160, loss: 0.044505681842565536
step: 170, loss: 0.07119208574295044
step: 180, loss: 0.015885518863797188
step: 190, loss: 0.004404056817293167
step: 200, loss: 0.005362851079553366
step: 210, loss: 0.08640384674072266
step: 220, loss: 0.04074140638113022
step: 230, loss: 0.04133203253149986
step: 240, loss: 0.004614831414073706
step: 250, loss: 0.10679202526807785
step: 260, loss: 0.05621139705181122
step: 270, loss: 0.006266979034990072
step: 280, loss: 0.18365998566150665
step: 290, loss: 0.002961982972919941
step: 300, loss: 0.021010875701904297
step: 310, loss: 0.0046910629607737064
step: 320, loss: 0.01832771487534046
step: 330, loss: 0.0050755771808326244
step: 340, loss: 0.003521141828969121
step: 350, loss: 0.0021629484836012125
step: 360, loss: 0.008704028092324734
step: 370, loss: 0.009451955556869507
step: 380, loss: 0.012189068831503391
step: 390, loss: 0.1430974006652832
step: 400, loss: 0.022416384890675545
step: 410, loss: 0.009126327000558376
step: 420, loss: 0.12020699679851532
step: 430, loss: 0.025855014100670815
step: 440, loss: 0.07562638819217682
step: 450, loss: 0.09876017272472382
step: 460, loss: 0.04172957316040993
```

```

step: 470, loss: 0.07716578245162964
step: 480, loss: 0.003174152225255966
step: 490, loss: 0.008468909189105034
step: 500, loss: 0.05130098760128021
step: 510, loss: 0.029395215213298798
step: 520, loss: 0.037778984755277634
step: 530, loss: 0.04430020600557327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9474662947466295, f1=0.9479502533394749,
best_f1=0.9479502533394749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.038427215069532394
step: 10, loss: 0.0019703300204128027
step: 20, loss: 0.025902563706040382
step: 30, loss: 0.0009482851601205766
step: 40, loss: 0.0025684963911771774
step: 50, loss: 0.007518215570598841
step: 60, loss: 0.003244789782911539
step: 70, loss: 0.0020314010325819254
step: 80, loss: 0.0008524390286765993
step: 90, loss: 0.03556136041879654
step: 100, loss: 0.004826240241527557
step: 110, loss: 0.011635750532150269
step: 120, loss: 0.06584350764751434
step: 130, loss: 0.0018681371584534645
step: 140, loss: 0.0007846925873309374
step: 150, loss: 0.0034811971709132195
step: 160, loss: 0.1439487338066101
step: 170, loss: 0.014941524714231491
step: 180, loss: 0.018607767298817635
step: 190, loss: 0.0635635033249855
step: 200, loss: 0.009147527627646923
step: 210, loss: 0.03446092829108238
step: 220, loss: 0.01685379631817341
step: 230, loss: 0.0482093021273613
step: 240, loss: 0.0033263086806982756
step: 250, loss: 0.057248037308454514
step: 260, loss: 0.0007467310060746968
step: 270, loss: 0.007428260985761881
step: 280, loss: 0.020618373528122902

```

```
step: 290, loss: 0.028581062331795692
step: 300, loss: 0.000391496520023793
step: 310, loss: 0.002768078353255987
step: 320, loss: 0.056272976100444794
step: 330, loss: 0.04384085536003113
step: 340, loss: 0.04227061569690704
step: 350, loss: 0.10276917368173599
step: 360, loss: 0.07598218321800232
step: 370, loss: 0.030809512361884117
step: 380, loss: 0.016952991485595703
step: 390, loss: 0.005173827055841684
step: 400, loss: 0.0014432963216677308
step: 410, loss: 0.009627067483961582
step: 420, loss: 0.0015881027793511748
step: 430, loss: 0.23492558300495148
step: 440, loss: 0.018501408398151398
step: 450, loss: 0.02526506967842579
step: 460, loss: 0.0019105508690699935
step: 470, loss: 0.052527766674757004
step: 480, loss: 0.25423678755760193
step: 490, loss: 0.023563610389828682
step: 500, loss: 0.011490504257380962
step: 510, loss: 0.07828742265701294
step: 520, loss: 0.08539528399705887
step: 530, loss: 0.05333017557859421
/usr/local/lib/python3.7/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=0.9468377635197067, f1=0.9447740757644911,
best_f1=0.9479502533394749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013365345075726509
step: 10, loss: 0.03973415121436119
step: 20, loss: 0.006755183916538954
step: 30, loss: 0.0014314415166154504
step: 40, loss: 0.026962468400597572
step: 50, loss: 0.001744818058796227
step: 60, loss: 0.16052035987377167
step: 70, loss: 0.03605787456035614
step: 80, loss: 0.023351281881332397
step: 90, loss: 0.0017357063479721546
step: 100, loss: 0.012245227582752705
step: 110, loss: 0.025400396436452866
step: 120, loss: 0.0033215240109711885
step: 130, loss: 0.00015643569349776953
step: 140, loss: 0.0002178760914830491
```

```

step: 150, loss: 0.0034844207111746073
step: 160, loss: 0.002553300466388464
step: 170, loss: 0.01807798258960247
step: 180, loss: 0.0009684594697318971
step: 190, loss: 0.009755411185324192
step: 200, loss: 0.0007518099155277014
step: 210, loss: 0.001318100607022643
step: 220, loss: 0.008238685317337513
step: 230, loss: 0.011227142065763474
step: 240, loss: 0.002131171990185976
step: 250, loss: 0.0007948413258418441
step: 260, loss: 0.013592411763966084
step: 270, loss: 0.0008488977327942848
step: 280, loss: 0.019727976992726326
step: 290, loss: 0.11442600190639496
step: 300, loss: 0.0017023893306031823
step: 310, loss: 0.019686011597514153
step: 320, loss: 0.028870537877082825
step: 330, loss: 0.011090272106230259
step: 340, loss: 0.002178780734539032
step: 350, loss: 0.004026546608656645
step: 360, loss: 0.0046342541463673115
step: 370, loss: 0.035670459270477295
step: 380, loss: 0.0004998401855118573
step: 390, loss: 0.0002463640703354031
step: 400, loss: 0.04049873352050781
step: 410, loss: 0.0006888293428346515
step: 420, loss: 0.002611772157251835
step: 430, loss: 0.0010414757998660207
step: 440, loss: 0.05444497987627983
step: 450, loss: 0.004550463054329157
step: 460, loss: 0.0009111876133829355
step: 470, loss: 0.006491406820714474
step: 480, loss: 0.03344017267227173
step: 490, loss: 0.0002901249099522829
step: 500, loss: 0.0041278344579041
step: 510, loss: 0.04769330844283104
step: 520, loss: 0.02133028395473957
step: 530, loss: 0.068163201212883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9507139567019807, f1=0.9463548830811555,
best_f1=0.9463548830811555

```

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

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

step: 0, loss: 0.003279436379671097

step: 10, loss: 0.00033632785198278725

step: 20, loss: 0.0836535319685936

step: 30, loss: 0.007873750291764736

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

step: 40, loss: 0.0230540931224823

step: 50, loss: 0.022920645773410797

step: 60, loss: 0.00026326000806875527

step: 70, loss: 0.019026560708880424

step: 80, loss: 0.001458174199797213

step: 90, loss: 0.0003001079603563994

step: 100, loss: 0.0011837175115942955

step: 110, loss: 0.01173827238380909

step: 120, loss: 0.0005759228370152414

step: 130, loss: 0.0007789114606566727

step: 140, loss: 0.021215636283159256

step: 150, loss: 0.002600093837827444

step: 160, loss: 0.003128377255052328

step: 170, loss: 0.002733407774940133

step: 180, loss: 0.0020619858987629414

step: 190, loss: 0.0012395946541801095

step: 200, loss: 0.00018740723317023367

step: 210, loss: 0.01777615025639534

step: 220, loss: 0.0017726451624184847

step: 230, loss: 0.0009156033047474921

step: 240, loss: 0.009935017675161362

step: 250, loss: 0.05727634206414223

step: 260, loss: 0.0003905218618456274

step: 270, loss: 0.1371704488992691

step: 280, loss: 0.1560097187757492

step: 290, loss: 0.001339217647910118

step: 300, loss: 0.0020053007174283266

step: 310, loss: 0.0030384757556021214

step: 320, loss: 0.0016252739587798715

step: 330, loss: 0.0006518086884170771

step: 340, loss: 0.08216825127601624

step: 350, loss: 0.002594933845102787

step: 360, loss: 0.044950343668460846

step: 370, loss: 0.0012574557913467288

step: 380, loss: 0.00025210267631337047

step: 390, loss: 0.00868627056479454

step: 400, loss: 0.00010038336768047884

step: 410, loss: 0.0019151894375681877

step: 420, loss: 0.000550043594557792

step: 430, loss: 0.0029783817008137703

```

step: 440, loss: 0.00048306575627066195
step: 450, loss: 0.0005602992605417967
step: 460, loss: 0.004042832180857658
step: 470, loss: 0.004836914129555225
step: 480, loss: 0.10138013958930969
step: 490, loss: 0.012670771218836308
step: 500, loss: 0.0015152256237342954
step: 510, loss: 0.002706402912735939
step: 520, loss: 0.08675684034824371
step: 530, loss: 0.004752345848828554
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9526000920386564, f1=0.9413936317489617,
best_f1=0.9413936317489617
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003870299260597676
step: 10, loss: 0.0012988222297281027
step: 20, loss: 0.0749763697385788
step: 30, loss: 0.0037824157625436783
step: 40, loss: 0.0005092684295959771
step: 50, loss: 0.006491462700068951
step: 60, loss: 0.00123923784121871
step: 70, loss: 0.007238345220685005
step: 80, loss: 0.03625241294503212
step: 90, loss: 0.0011340141063556075
step: 100, loss: 0.00011811127478722483
step: 110, loss: 0.00012035747931804508
step: 120, loss: 0.0016734374221414328
step: 130, loss: 0.018212974071502686
step: 140, loss: 7.682014256715775e-05
step: 150, loss: 0.013752225786447525
step: 160, loss: 0.00028552181902341545
step: 170, loss: 9.187018440570682e-05
step: 180, loss: 0.0008742845384404063
step: 190, loss: 0.001663109054788947
step: 200, loss: 0.0006313152844086289
step: 210, loss: 0.00856149010360241
step: 220, loss: 0.0010747014312073588
step: 230, loss: 0.048347849398851395
step: 240, loss: 0.0009370602201670408
step: 250, loss: 0.0009183272486552596

```

```

step: 260, loss: 8.405202970607206e-05
step: 270, loss: 0.0033539545256644487
step: 280, loss: 0.000686576240696013
step: 290, loss: 0.0004441042256075889
step: 300, loss: 0.0014394995523616672
step: 310, loss: 0.0009687396232038736
step: 320, loss: 0.0015841034473851323
step: 330, loss: 0.0020520102698355913
step: 340, loss: 0.007058931980282068
step: 350, loss: 0.0008800961659289896
step: 360, loss: 0.0036184084601700306
step: 370, loss: 0.000298507628031075
step: 380, loss: 0.007002009078860283
step: 390, loss: 0.00042336180922575295
step: 400, loss: 0.00039948022458702326
step: 410, loss: 0.0029811798594892025
step: 420, loss: 0.000352948612999171
step: 430, loss: 0.0001940239017130807
step: 440, loss: 0.000275657104793936
step: 450, loss: 0.0005158032290637493
step: 460, loss: 0.001685841241851449
step: 470, loss: 0.009470013901591301
step: 480, loss: 0.17962518334388733
step: 490, loss: 0.00035469612339511514
step: 500, loss: 0.004832111299037933
step: 510, loss: 0.003187387017533183
step: 520, loss: 0.001978836487978697
step: 530, loss: 0.03820968046784401
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9546716003700277, f1=0.9442653155228006,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005658808164298534
step: 10, loss: 0.021897664293646812
step: 20, loss: 0.011139343492686749
step: 30, loss: 0.027224473655223846
step: 40, loss: 5.2755338401766494e-05
step: 50, loss: 0.0026575876399874687
step: 60, loss: 0.0010344400070607662
step: 70, loss: 9.82924466370605e-05

```


step: 80, loss: 0.003774086944758892
step: 90, loss: 0.0006198359187692404
step: 100, loss: 0.0002611312665976584
step: 110, loss: 0.00017751124687492847
step: 120, loss: 0.003933883272111416
step: 130, loss: 0.0012657204642891884
step: 140, loss: 0.003918319940567017
step: 150, loss: 0.20181767642498016
step: 160, loss: 0.0008084605797193944
step: 170, loss: 0.06627834588289261
step: 180, loss: 0.00033310684375464916
step: 190, loss: 0.0033973536919802427
step: 200, loss: 0.00333578628487885
step: 210, loss: 0.012687042355537415
step: 220, loss: 0.0058774277567863464
step: 230, loss: 0.00011726676893886179
step: 240, loss: 0.0005794314783997834
step: 250, loss: 0.00010249250044580549
step: 260, loss: 7.890919368946925e-05
step: 270, loss: 0.002843594178557396
step: 280, loss: 0.008442776277661324
step: 290, loss: 0.0013883357169106603
step: 300, loss: 0.000579953077249229
step: 310, loss: 0.0006816885434091091
step: 320, loss: 0.00037668869481422007
step: 330, loss: 0.014854753389954567
step: 340, loss: 9.656642214395106e-05
step: 350, loss: 8.363516826648265e-05
step: 360, loss: 0.0007009709952399135
step: 370, loss: 0.0003591321292333305
step: 380, loss: 0.0011461017420515418
step: 390, loss: 0.12927012145519257
step: 400, loss: 0.0002632486284710467
step: 410, loss: 0.008400114253163338
step: 420, loss: 0.0027075642719864845
step: 430, loss: 0.04483721777796745
step: 440, loss: 0.008621301501989365
step: 450, loss: 0.0009637451730668545
step: 460, loss: 0.0009283084655180573
step: 470, loss: 0.0003536526928655803
step: 480, loss: 0.0009443160961382091
step: 490, loss: 0.0027369579765945673
step: 500, loss: 0.0380672886967659
step: 510, loss: 0.001163980457931757
step: 520, loss: 0.0029221167787909508
step: 530, loss: 0.0005945426528342068
/usr/local/lib/python3.7/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=0.951851851851852, f1=0.9476609541454377,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008189060725271702
step: 10, loss: 0.009763339534401894
step: 20, loss: 0.00037165722460485995
step: 30, loss: 0.03738235682249069
step: 40, loss: 0.00024429880431853235
step: 50, loss: 0.0004625660367310047
step: 60, loss: 0.00027879985282197595
step: 70, loss: 0.00031649612355977297
step: 80, loss: 0.09028026461601257
step: 90, loss: 0.009117143228650093
step: 100, loss: 0.0024176996666938066
step: 110, loss: 0.00035053450847044587
step: 120, loss: 0.0004077472258359194
step: 130, loss: 0.0011299545876681805
step: 140, loss: 0.0001239019475178793
step: 150, loss: 0.00020737957675009966
step: 160, loss: 0.0007004768121987581
step: 170, loss: 0.014893777668476105
step: 180, loss: 0.0002825551200658083
step: 190, loss: 0.0026766322553157806
step: 200, loss: 0.00032957596704363823
step: 210, loss: 0.0002580684667918831
step: 220, loss: 0.040250737220048904
step: 230, loss: 0.022960757836699486
step: 240, loss: 0.00020397987100295722
step: 250, loss: 0.005612972658127546
step: 260, loss: 0.03257812559604645
step: 270, loss: 0.0002685926156118512
step: 280, loss: 0.00408347649499774
step: 290, loss: 0.035702869296073914
step: 300, loss: 0.00021637495956383646
step: 310, loss: 0.02256256528198719
step: 320, loss: 0.0007523070671595633
step: 330, loss: 0.005192930810153484
step: 340, loss: 0.00026269257068634033
step: 350, loss: 0.004717694595456123
step: 360, loss: 0.0002804447431117296
step: 370, loss: 0.007274976000189781
step: 380, loss: 0.00014263046614360064
step: 390, loss: 0.0003873464302159846
step: 400, loss: 0.001678009401075542
step: 410, loss: 0.0003698798827826977
```

```

step: 420, loss: 5.748546027461998e-05
step: 430, loss: 7.855273724999279e-05
step: 440, loss: 0.002316083526238799
step: 450, loss: 6.797997775720432e-05
step: 460, loss: 9.747096919454634e-05
step: 470, loss: 0.00014188756176736206
step: 480, loss: 0.003219504142180085
step: 490, loss: 0.003256527939811349
step: 500, loss: 0.0011348684784024954
step: 510, loss: 0.004295120015740395
step: 520, loss: 0.00012531709217000753
step: 530, loss: 0.0002165243640774861
/usr/local/lib/python3.7/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=0.9457579972183587, f1=0.9454209065679925,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006694335024803877
step: 10, loss: 0.00017337771714664996
step: 20, loss: 8.455953502561897e-05
step: 30, loss: 4.134103801334277e-05
step: 40, loss: 0.00011766474926844239
step: 50, loss: 0.0007133916369639337
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 60, loss: 0.00018301300588063896
step: 70, loss: 8.957442332757637e-05
step: 80, loss: 7.076676411088556e-05
step: 90, loss: 0.002396013354882598
step: 100, loss: 0.0026254327967762947
step: 110, loss: 0.00042435614159330726
step: 120, loss: 0.07151860743761063
step: 130, loss: 0.0034934845753014088
step: 140, loss: 0.01399106252938509
step: 150, loss: 0.0014330418780446053
step: 160, loss: 0.00022397008433472365
step: 170, loss: 0.002675052499398589
step: 180, loss: 0.0016160743543878198
step: 190, loss: 0.0012553947744891047
step: 200, loss: 0.0024388532619923353
step: 210, loss: 0.0014193904353305697
step: 220, loss: 0.0040984139777719975
step: 230, loss: 0.0020457329228520393
step: 240, loss: 0.00019796894048340619
step: 250, loss: 7.49566315789707e-05
step: 260, loss: 0.0025112521834671497

```

```

step: 270, loss: 0.054770272225141525
step: 280, loss: 0.027420571073889732
step: 290, loss: 0.00016806733037810773
step: 300, loss: 3.717348590726033e-05
step: 310, loss: 0.0001430190313840285
step: 320, loss: 0.003633533837273717
step: 330, loss: 0.00015500764129683375
step: 340, loss: 0.00923953391611576
step: 350, loss: 0.0007356277201324701
step: 360, loss: 0.0001475805911468342
step: 370, loss: 0.003285285783931613
step: 380, loss: 0.03852701932191849
step: 390, loss: 0.006879650522023439
step: 400, loss: 0.002631891518831253
step: 410, loss: 0.002177928574383259
step: 420, loss: 0.0005670522805303335
step: 430, loss: 0.00057653320254758
step: 440, loss: 0.006868514697998762
step: 450, loss: 5.247913577477448e-05
step: 460, loss: 0.01566944271326065
step: 470, loss: 0.00033230037661269307
step: 480, loss: 6.052336539141834e-05
step: 490, loss: 0.0002526686294004321
step: 500, loss: 0.013468008488416672
step: 510, loss: 0.00016064821102190763
step: 520, loss: 0.0006033882382325828
step: 530, loss: 0.00046215479960665107
/usr/local/lib/python3.7/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=0.9465437788018433, f1=0.9452369995398067,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017780194757506251
step: 10, loss: 0.0029377401806414127
step: 20, loss: 4.535835614660755e-05
step: 30, loss: 0.00027148809749633074
step: 40, loss: 6.713486072840169e-05
step: 50, loss: 0.00013684768055099994
step: 60, loss: 0.025853214785456657
step: 70, loss: 7.81788767199032e-05
step: 80, loss: 0.0007577387150377035
step: 90, loss: 4.51409250672441e-05
step: 100, loss: 0.013198803178966045
step: 110, loss: 0.0009638387709856033
step: 120, loss: 7.062801159918308e-05

```

```

step: 130, loss: 2.039933497144375e-05
step: 140, loss: 0.013166535645723343
step: 150, loss: 2.343904270674102e-05
step: 160, loss: 3.761234620469622e-05
step: 170, loss: 3.39543548761867e-05
step: 180, loss: 9.198443876812235e-05
step: 190, loss: 0.0031952359713613987
step: 200, loss: 0.006814274471253157
step: 210, loss: 7.005036604823545e-05
step: 220, loss: 0.0016832828987389803
step: 230, loss: 7.229892071336508e-05
step: 240, loss: 0.00045501903514377773
step: 250, loss: 1.765756678651087e-05
step: 260, loss: 1.7106331142713316e-05
step: 270, loss: 0.02995743602514267
step: 280, loss: 0.00043513061245903373
step: 290, loss: 0.025693807750940323
step: 300, loss: 0.0018231681315228343
step: 310, loss: 4.1091425373451784e-05
step: 320, loss: 4.281777364667505e-05
step: 330, loss: 6.832607323303819e-05
step: 340, loss: 0.0022991166915744543
step: 350, loss: 4.0231378079624847e-05
step: 360, loss: 0.0047817290760576725
step: 370, loss: 0.003051621839404106
step: 380, loss: 5.788333874079399e-05
step: 390, loss: 0.000357926357537508
step: 400, loss: 0.0002800285874400288
step: 410, loss: 0.0023121859412640333
step: 420, loss: 0.002843467751517892
step: 430, loss: 1.666674688749481e-05
step: 440, loss: 0.0002136266848538071
step: 450, loss: 0.00024154692073352635
step: 460, loss: 0.16480752825737
step: 470, loss: 0.0003555468574631959
step: 480, loss: 0.0007034083246253431
step: 490, loss: 0.0007770408992655575
step: 500, loss: 0.0009067401988431811
step: 510, loss: 0.000748734746593982
step: 520, loss: 2.020565989369061e-05
step: 530, loss: 3.262455720687285e-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=0.9465020576131687, f1=0.9457858769931662,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.794591586687602e-05
step: 10, loss: 2.493274223525077e-05
step: 20, loss: 7.790896779624745e-05
step: 30, loss: 0.001286600949242711
step: 40, loss: 0.0008371178992092609
step: 50, loss: 0.022827167063951492
step: 60, loss: 0.00038188035250641406
step: 70, loss: 0.0040501439943909645
step: 80, loss: 0.0001348528021480888
step: 90, loss: 3.699793160194531e-05
step: 100, loss: 0.0009768889285624027
step: 110, loss: 0.00980242807418108
step: 120, loss: 0.00011329713015584275
step: 130, loss: 0.0007963977986946702
step: 140, loss: 3.2783555070636794e-05
step: 150, loss: 7.099381764419377e-05
step: 160, loss: 3.335422297823243e-05
step: 170, loss: 9.250208677258343e-05
step: 180, loss: 0.002942355815321207
step: 190, loss: 0.0007098325295373797
step: 200, loss: 0.00014119147090241313
step: 210, loss: 3.400146306375973e-05
step: 220, loss: 0.002204711316153407
step: 230, loss: 4.0853126847650856e-05
step: 240, loss: 0.00013668044994119555
step: 250, loss: 2.8709584512398578e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 260, loss: 5.1439365051919594e-05
step: 270, loss: 0.0038733785040676594
step: 280, loss: 0.00124463124666363
step: 290, loss: 0.0015924472827464342
step: 300, loss: 0.005540471989661455
step: 310, loss: 0.00014895029016770422
step: 320, loss: 0.0001294969697482884
step: 330, loss: 6.231042789295316e-05
step: 340, loss: 6.95881899446249e-05
step: 350, loss: 0.0004529731231741607
step: 360, loss: 0.006146138999611139
step: 370, loss: 0.0011841091327369213
step: 380, loss: 0.00031259877141565084
step: 390, loss: 0.0002275742735946551
step: 400, loss: 4.6284112613648176e-05
step: 410, loss: 0.0019660783000290394
step: 420, loss: 0.004271214362233877
step: 430, loss: 0.03611451014876366
step: 440, loss: 0.00011473353515611961
step: 450, loss: 0.00010371697862865403
```

```

step: 460, loss: 3.691844176501036e-05
step: 470, loss: 0.03399944677948952
step: 480, loss: 0.0013890694826841354
step: 490, loss: 0.0017647091299295425
step: 500, loss: 0.00034981098724529147
step: 510, loss: 0.0012528377119451761
step: 520, loss: 7.408553210552782e-05
step: 530, loss: 0.00017598214617464691
/usr/local/lib/python3.7/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=0.9494199535962878, f1=0.9445983379501385,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00244833342730999
step: 10, loss: 0.00017872195167001337
step: 20, loss: 0.0002659836027305573
step: 30, loss: 0.000497701286803931
step: 40, loss: 0.0014191092923283577
step: 50, loss: 0.003743582870811224
step: 60, loss: 0.00014293214189819992
step: 70, loss: 9.331112232757732e-05
step: 80, loss: 7.001654012128711e-05
step: 90, loss: 0.017225516960024834
step: 100, loss: 0.000617884797975421
step: 110, loss: 0.0003086248179897666
step: 120, loss: 0.00010166665015276521
step: 130, loss: 2.7454418159322813e-05
step: 140, loss: 0.0005155659164302051
step: 150, loss: 0.00025013196864165366
step: 160, loss: 0.0013422559713944793
step: 170, loss: 5.567770131165162e-05
step: 180, loss: 3.8308971852529794e-05
step: 190, loss: 7.480855856556445e-05
step: 200, loss: 3.196121906512417e-05
step: 210, loss: 0.0010733790695667267
step: 220, loss: 5.469773896038532e-05
step: 230, loss: 0.004370496608316898
step: 240, loss: 0.0002963678853120655
step: 250, loss: 8.286194497486576e-05
step: 260, loss: 0.012539583258330822
step: 270, loss: 3.419976565055549e-05
step: 280, loss: 4.558312866720371e-05
step: 290, loss: 9.890373621601611e-05
step: 300, loss: 4.909470590064302e-05
step: 310, loss: 0.0005922533455304801

```

```

step: 320, loss: 0.003523283638060093
step: 330, loss: 0.00838597770780325
step: 340, loss: 0.0024711082223802805
step: 350, loss: 0.00042142526945099235
step: 360, loss: 0.0001017119939206168
step: 370, loss: 0.00016277329996228218
step: 380, loss: 5.337402762961574e-05
step: 390, loss: 0.004480015952140093
step: 400, loss: 0.00021561172616202384
step: 410, loss: 0.001469236216507852
step: 420, loss: 6.987246160861105e-05
step: 430, loss: 0.0008294142317026854
step: 440, loss: 0.006351996213197708
step: 450, loss: 0.00011768508556997404
step: 460, loss: 0.0014531435444951057
step: 470, loss: 0.00011099074617959559
step: 480, loss: 0.0003662262170109898
step: 490, loss: 0.0010357317514717579
step: 500, loss: 5.500225961441174e-05
step: 510, loss: 3.73999573639594e-05
step: 520, loss: 0.0012329345336183906
step: 530, loss: 0.0001148536684922874
/usr/local/lib/python3.7/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=0.9491841491841492, f1=0.945707656612529,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027797953225672245
step: 10, loss: 9.864776802714914e-05
step: 20, loss: 0.020863058045506477
step: 30, loss: 3.640868089860305e-05
step: 40, loss: 0.0007500179926864803
step: 50, loss: 0.000320190389174968
step: 60, loss: 2.442213735776022e-05
step: 70, loss: 0.00045753346057608724
step: 80, loss: 2.022394437517505e-05
step: 90, loss: 0.00020473172480706125
step: 100, loss: 0.007514892145991325
step: 110, loss: 0.000152128966874443
step: 120, loss: 0.00011501603876240551
step: 130, loss: 0.006134508177638054
step: 140, loss: 0.0009902139427140355
step: 150, loss: 9.428375051356852e-05
step: 160, loss: 0.00017892431060317904
step: 170, loss: 0.0007270397618412971

```



```

step: 180, loss: 0.00027524889446794987
step: 190, loss: 0.0003221613878849894
step: 200, loss: 0.0017481619725003839
step: 210, loss: 3.495573764666915e-05
step: 220, loss: 7.942438242025673e-05
step: 230, loss: 0.003025301732122898
step: 240, loss: 3.322802876937203e-05
step: 250, loss: 0.00013187187141738832
step: 260, loss: 0.0013410556130111217
step: 270, loss: 0.00023140232951845974
step: 280, loss: 7.711437501711771e-05
step: 290, loss: 0.00015411715139634907
step: 300, loss: 0.0004400164762046188
step: 310, loss: 0.0641755685210228
step: 320, loss: 0.00035978262894786894
step: 330, loss: 0.0009529992239549756
step: 340, loss: 0.0007723110029473901
step: 350, loss: 0.0004533551400527358
step: 360, loss: 0.0027506283950060606
step: 370, loss: 9.837311517912894e-05
step: 380, loss: 0.0018011979991570115
step: 390, loss: 0.028367826715111732
step: 400, loss: 0.0008186884224414825
step: 410, loss: 5.50663608009927e-05
step: 420, loss: 0.00016965022950898856
step: 430, loss: 8.43207526486367e-05
step: 440, loss: 0.0036829132586717606
step: 450, loss: 0.0002415235503576696
step: 460, loss: 0.00013788104115519673
step: 470, loss: 0.0005414828774519265
step: 480, loss: 2.6117055313079618e-05
step: 490, loss: 6.173216388560832e-05
step: 500, loss: 6.055623816791922e-05
step: 510, loss: 0.003826538100838661
step: 520, loss: 0.0002468927123118192
step: 530, loss: 7.322304736590013e-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=0.9498141263940519, f1=0.9434137291280148,
best_f1=0.9442653155228006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.8333353081252426e-05
step: 10, loss: 3.11507101287134e-05
step: 20, loss: 0.00011418702342780307
step: 30, loss: 0.013676309026777744

```

step: 40, loss: 0.04636211320757866
step: 50, loss: 0.016719287261366844
step: 60, loss: 0.00012258645438123494
step: 70, loss: 0.002644267166033387
step: 80, loss: 0.0001260609715245664
step: 90, loss: 0.00015034472744446248
step: 100, loss: 8.477118535665795e-05
step: 110, loss: 0.001730044954456389
step: 120, loss: 0.004272202495485544
step: 130, loss: 0.14110572636127472
step: 140, loss: 9.252206655219197e-05
step: 150, loss: 3.0143139156280085e-05
step: 160, loss: 0.0001393366837874055
step: 170, loss: 3.5688484786078334e-05
step: 180, loss: 4.4404070649761707e-05
step: 190, loss: 0.0002864942071028054
step: 200, loss: 0.0006705565610900521
step: 210, loss: 0.0009385317680425942
step: 220, loss: 0.00012909444922115654
step: 230, loss: 8.409828296862543e-05
step: 240, loss: 0.003145151538774371
step: 250, loss: 8.628707291791216e-05
step: 260, loss: 0.0003830782079603523
step: 270, loss: 1.5426212485181168e-05
step: 280, loss: 4.346125933807343e-05
step: 290, loss: 2.3021742890705355e-05
step: 300, loss: 3.2225503673544154e-05
step: 310, loss: 0.00011560526036191732
step: 320, loss: 0.00043163594091311097
step: 330, loss: 3.3834210626082495e-05
step: 340, loss: 0.00022913393331691623
step: 350, loss: 0.01222384162247181
step: 360, loss: 0.0021045173052698374
step: 370, loss: 5.136957406648435e-05
step: 380, loss: 0.00023598178813699633
step: 390, loss: 0.032554347068071365
step: 400, loss: 6.81009260006249e-05
step: 410, loss: 0.00011648531653918326
step: 420, loss: 6.874270911794156e-05
step: 430, loss: 0.00019729437190108
step: 440, loss: 0.00019141445227432996
step: 450, loss: 3.892857057508081e-05
step: 460, loss: 4.360782622825354e-05
step: 470, loss: 0.0027236088644713163
step: 480, loss: 6.251510058064014e-05
step: 490, loss: 1.3641882105730474e-05
step: 500, loss: 0.0004835189611185342
step: 510, loss: 8.95545381354168e-05

```

step: 520, loss: 4.7194207581924275e-05
step: 530, loss: 0.0005498392274603248
/usr/local/lib/python3.7/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=0.9500462534690102, f1=0.943239501615136,
best_f1=0.9442653155228006
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 [00:28, 201.19it/s]
load_f1 = 0.952513966480447
real_f1 = 0.9514018691588785
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 196.48it/s]

```