

HLow_10_1_2_bert

August 13, 2022

0.1 Install packages

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 23.06 MiB/s, done.
Resolving deltas: 100% (142/142), done.
Checking out files: 100% (128/128), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 10.0 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 47.3 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 72.7 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 56.9 MB/s
Requirement already satisfied: sklearn==0.0 in
```

```

/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
  | 6.4 MB 14.2 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 21.71 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 6.7 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 55.4 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
  | 753.4 MB 7.0 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
  | 56 kB 6.3 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
  | 884 kB 64.7 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 59.5 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: 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: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Collecting catalogue<1.1.0,>=0.0.7
    Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting blis<0.5.0,>=0.4.0
    Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
      |                               | 3.7 MB 51.3 MB/s
Collecting srsly<1.1.0,>=0.1.0
    Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
      |                               | 184 kB 72.3 MB/s
Requirement already satisfied: setuptools in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (57.4.0)
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 44.4 MB/s
Collecting plac<1.2.0,>=0.9.6
    Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
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: filelock in /usr/local/lib/python3.7/dist-

```

```

packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    |                               | 880 kB 51.7 MB/s
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    |                               | 3.0 MB 50.8 MB/s
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: 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: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: 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=33738a04b99f82be36e7f453dc8fc24d9439efdd0154061ab1e67ff797911d37
  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=228ad4fba13841a5a204e3a47a96387bf8b6ddac266ff114bdc00d381b22d403
Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
Attempting uninstall: numpy
  Found existing installation: numpy 1.21.6
  Uninstalling numpy-1.21.6:
    Successfully uninstalled numpy-1.21.6
Attempting uninstall: tqdm
  Found existing installation: tqdm 4.64.0
  Uninstalling tqdm-4.64.0:
    Successfully uninstalled tqdm-4.64.0
Attempting uninstall: srsly
  Found existing installation: srsly 2.4.4
  Uninstalling srsly-2.4.4:
    Successfully uninstalled srsly-2.4.4
Attempting uninstall: regex
  Found existing installation: regex 2022.6.2
  Uninstalling regex-2022.6.2:
    Successfully uninstalled regex-2022.6.2
Attempting uninstall: blis
  Found existing installation: blis 0.7.8
  Uninstalling blis-0.7.8:
    Successfully uninstalled blis-0.7.8
Attempting uninstall: thinc
  Found existing installation: thinc 8.1.0
  Uninstalling thinc-8.1.0:
    Successfully uninstalled thinc-8.1.0
Attempting uninstall: catalogue
  Found existing installation: catalogue 2.0.8
  Uninstalling catalogue-2.0.8:
    Successfully uninstalled catalogue-2.0.8
Attempting uninstall: torch
  Found existing installation: torch 1.12.0+cu113
  Uninstalling torch-1.12.0+cu113:
    Successfully uninstalled torch-1.12.0+cu113
Attempting uninstall: spacy
  Found existing installation: spacy 3.4.1
  Uninstalling spacy-3.4.1:
    Successfully uninstalled spacy-3.4.1
Attempting uninstall: nltk
  Found existing installation: nltk 3.7

```

```

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

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

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

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

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

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

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

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

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

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

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

```

```
[1]: True
```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10117, done.
remote: Counting objects: 100% (233/233), done.
remote: Compressing objects: 100% (165/165), done.
remote: Total 10117 (delta 119), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.97 MiB | 23.99 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-jkww97t6
Created temporary directory: /tmp/pip-req-tracker-5aelxtds
Initialized build tracking at /tmp/pip-req-tracker-5aelxtds
Created build tracker: /tmp/pip-req-tracker-5aelxtds
Entered build tracker: /tmp/pip-req-tracker-5aelxtds
```

```

Created temporary directory: /tmp/pip-install-tfd7ogmx
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-_nw0tf2m
  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-5aelxts'
  Running setup.py (path:/tmp/pip-req-build-_nw0tf2m/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-uhlc3xbe
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-uhlc3xbe/apex.egg-info
  writing /tmp/pip-pip-egg-info-uhlc3xbe/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-uhlc3xbe/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-uhlc3xbe/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-uhlc3xbe/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-uhlc3xbe/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-_nw0tf2m 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-5aelxts'
Created temporary directory: /tmp/pip-unpack-15ak0yuz
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-mt80mmb5
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-mt80mmb5
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-_nw0tf2m/setup.py''';

```



```
__file__='''/tmp/pip-req-build-nw0tf2m/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-mt80mmb5
```

```
torch.__version__ = 1.4.0
```

```
running bdist_wheel
running build
running build_py
creating build
creating build/lib
creating build/lib/apex
copying apex/__init__.py -> build/lib/apex
copying apex/_autocast_utils.py -> build/lib/apex
creating build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
creating build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
creating build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/__init__.py ->
build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/multi_tensor_apply.py ->
build/lib/apex/multi_tensor_apply
creating build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
creating build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
```

```

copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
creating build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
copying apex/normalization/__init__.py -> build/lib/apex/normalization
creating build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
creating build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
creating build/lib/apex/transformer/layers
copying apex/transformer/layers/layer_norm.py ->

```

```

build/lib/apex/transformer/layers
  copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
  creating build/lib/apex/transformer/_data
  copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
  copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
  creating build/lib/apex/transformer/testing
  copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing
  creating build/lib/apex/transformer/functional
  copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
  copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
  creating build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/mappings.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/layers.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
  creating build/lib/apex/transformer/amp
  copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp

```

```

    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py
-> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/__init__.py ->
build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/lib/apex/contrib/conv_bias_relu
    creating build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm
    creating build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->

```

```

build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/__init__.py ->
build/lib/apex/contrib/multihead_attn
  copying
apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
  creating build/lib/apex/contrib/xentropy
  copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
  copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
  creating build/lib/apex/contrib/sparsity
  copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
  copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
  copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
  copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
  creating build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
  creating build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/transducer.py ->

```

```

build/lib/apex/contrib/transducer
  creating build/lib/apex/contrib/index_mul_2d
  copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
  copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
  creating build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
  creating build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
  creating build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
  creating build/lib/apex/amp/lists
  copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
  copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
  copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
  copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
  installing to build/bdist.linux-x86_64/wheel
  running install
  running install_lib
  creating build/bdist.linux-x86_64
  creating build/bdist.linux-x86_64/wheel
  creating build/bdist.linux-x86_64/wheel/apex
  creating build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel

```

```

    copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py
-> build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_without_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel

```

```

    copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/distributed_test_base.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/microbatches.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/memory.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/layers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/cross_entropy.py ->

```



```

build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/amp/grad_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  creating build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/test.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/bottleneck_module_test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
  creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  copying build/lib/apex/contrib/layer_norm/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn

```

```

    copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers

```

```

    creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchange_module_tests.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/wrap.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/rnn_compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_initialize.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_amp_state.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_process_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp

```

```

    copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
    creating build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    running install_egg_info
    running egg_info
    creating apex.egg-info
    writing apex.egg-info/PKG-INFO
    writing dependency_links to apex.egg-info/dependency_links.txt
    writing top-level names to apex.egg-info/top_level.txt
    writing manifest file 'apex.egg-info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file 'apex.egg-info/SOURCES.txt'
    Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
    running install_scripts
    adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
    creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
    creating '/tmp/pip-wheel-mt80mb5/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=7ad7e3487bdab4a9838fdb4eead41d78a2965379cb6d9429d695462e8c5c1bcf
Stored in directory: /tmp/pip-ephem-wheel-cache-
jkwk97t6/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-5aelxtds'
/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 6.9 MB/s
Collecting awscli
  Downloading awscli-1.25.49-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 58.4 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting botocore==1.27.49
  Downloading botocore-1.27.49-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 51.6 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 62.9 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 4.5 MB/s
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.49->awscli) (2.8.2)
Collecting jmespath<2.0.0,>=0.7.1
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.49->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
  Attempting uninstall: urllib3
    Found existing installation: urllib3 1.24.3
    Uninstalling urllib3-1.24.3:
      Successfully uninstalled urllib3-1.24.3
  Attempting uninstall: rsa
    Found existing installation: rsa 4.9
    Uninstalling rsa-4.9:
      Successfully uninstalled rsa-4.9
  Attempting uninstall: docutils
    Found existing installation: docutils 0.17.1
    Uninstalling docutils-0.17.1:
      Successfully uninstalled docutils-0.17.1
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.
panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is
incompatible.
fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is
incompatible.
Successfully installed awscli-1.25.49 botocore-1.27.49 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```

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

```

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

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 1054, done.
remote: Counting objects: 100% (124/124), done.
remote: Compressing objects: 100% (91/91), done.
remote: Total 1054 (delta 51), reused 46 (delta 21), pack-reused 930
Receiving objects: 100% (1054/1054), 257.86 MiB | 22.98 MiB/s, done.
Resolving deltas: 100% (635/635), done.
Checking out files: 100% (1304/1304), done.

```

0.4 remove e move

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

```

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

```

```
[7]: !rm configs.json
!mv -n /content/ditto/IC/datasasErros/HLow_10_1_2/configs.json /content/ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

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

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

Downloading: 100% 433/433 [00:00<00:00, 593kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 19.2MB/s]

Downloading: 100% 440M/440M [00:10<00:00, 43.8MB/s]

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

Defaults for this optimization level are:

enabled : True

opt_level : 02

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5531837940216064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45454545454545453, f1=0.45161290322580644,
best_f1=0.45161290322580644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5068519711494446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.4736842105263159, f1=0.45, best_f1=0.45
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4841437339782715
/usr/local/lib/python3.7/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=0.4736842105263159, f1=0.4615384615384615, best_f1=0.45
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2573987543582916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7096774193548386, 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: 0.16088981926441193
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7272727272727273, f1=0.6842105263157894,
best_f1=0.6842105263157894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11303137987852097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8125000000000001, f1=0.742857142857143,
best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0174476969987154
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7647058823529412,
best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0076437159441411495
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.8, best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032873516902327538
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.75, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010823794640600681
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.7878787878787878, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004068554844707251
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.7878787878787878, best_f1=0.75

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005063862539827824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7999999999999999, f1=0.8125000000000001, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004686933942139149
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.8125000000000001, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021400518715381622
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.8125000000000001, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008363359607756138
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.8125000000000001, best_f1=0.75
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, 122099.06it/s]
/usr/local/lib/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 = 0.8484848484848484
real_f1 = 0.8484848484848484
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 [00:15, 274.92it/s]

```

1.1.2 DBLP-ACM - Running the matcher

```

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

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

# Step 3: run the matcher

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

```
step: 0, loss: 0.6306656002998352
```

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

```

step: 10, loss: 0.5773899555206299
step: 20, loss: 0.24668143689632416
step: 30, loss: 0.10147152841091156
step: 40, loss: 0.17137646675109863
step: 50, loss: 0.03738733381032944
step: 60, loss: 0.03550444543361664
step: 70, loss: 0.045998699963092804
step: 80, loss: 0.07862647622823715
step: 90, loss: 0.18613523244857788
step: 100, loss: 0.01906968094408512
step: 110, loss: 0.04749462381005287
step: 120, loss: 0.009648052975535393
step: 130, loss: 0.009322845377027988
step: 140, loss: 0.002872493350878358
step: 150, loss: 0.005915128625929356
step: 160, loss: 0.002510810736566782
step: 170, loss: 0.14370566606521606
step: 180, loss: 0.0431482270359993
step: 190, loss: 0.0033566290512681007
step: 200, loss: 0.07395268231630325
step: 210, loss: 0.005354343913495541
step: 220, loss: 0.05526697263121605
step: 230, loss: 0.02014361321926117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9909502262443439, f1=0.9829738933030647,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014620448928326368
step: 10, loss: 0.010902822948992252
step: 20, loss: 0.02049606293439865
step: 30, loss: 0.19145815074443817
step: 40, loss: 0.01955045387148857
step: 50, loss: 0.003426674986258149
step: 60, loss: 0.0018314735498279333
step: 70, loss: 0.21617071330547333
step: 80, loss: 0.0018053136300295591
step: 90, loss: 0.012683629989624023
step: 100, loss: 0.07621658593416214
step: 110, loss: 0.045236434787511826
step: 120, loss: 0.0015387794701382518

```

```

step: 130, loss: 0.0082610584795475
step: 140, loss: 0.0027688199188560247
step: 150, loss: 0.025737343356013298
step: 160, loss: 0.004401934798806906
step: 170, loss: 0.002482939511537552
step: 180, loss: 0.0043367380276322365
step: 190, loss: 0.003397288965061307
step: 200, loss: 0.005052558612078428
step: 210, loss: 0.0008025270653888583
step: 220, loss: 0.08534470945596695
step: 230, loss: 0.07704003155231476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932584269662922, f1=0.9876265466816648,
best_f1=0.9876265466816648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028721645940095186
step: 10, loss: 0.0008029064629226923
step: 20, loss: 0.0005330194835551083
step: 30, loss: 0.002936652163043618
step: 40, loss: 0.14962217211723328
step: 50, loss: 0.002667322987690568
step: 60, loss: 0.0010005198419094086
step: 70, loss: 0.0010327775962650776
step: 80, loss: 0.0003560180775821209
step: 90, loss: 0.07439574599266052
step: 100, loss: 0.0015114994021132588
step: 110, loss: 0.0004208233440294862
step: 120, loss: 0.01298461202532053
step: 130, loss: 0.0011257831938564777
step: 140, loss: 0.0009117741719819605
step: 150, loss: 0.06623538583517075
step: 160, loss: 0.04125344008207321
step: 170, loss: 0.0019330023787915707
step: 180, loss: 0.00936009269207716
step: 190, loss: 0.003060175571590662
step: 200, loss: 0.0018708414863795042
step: 210, loss: 0.0009344271384179592
step: 220, loss: 0.0008188950596377254
step: 230, loss: 0.00021485929028131068
/usr/local/lib/python3.7/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=0.9932432432432432, f1=0.987598647125141,
best_f1=0.9876265466816648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026051816530525684
step: 10, loss: 0.00022220650862436742
step: 20, loss: 0.0003894380643032491
step: 30, loss: 0.00029687577625736594
step: 40, loss: 0.0007859916076995432
step: 50, loss: 0.0005848389118909836
step: 60, loss: 0.0010435138829052448
step: 70, loss: 0.00020517990924417973
step: 80, loss: 0.007948153652250767
step: 90, loss: 0.0070052314549684525
step: 100, loss: 0.0001929295831359923
step: 110, loss: 0.0001099399451049976
step: 120, loss: 0.028715908527374268
step: 130, loss: 0.0017270686803385615
step: 140, loss: 0.00011328642722219229
step: 150, loss: 0.04740417003631592
step: 160, loss: 0.00020019781368318945
step: 170, loss: 0.0010057664476335049
step: 180, loss: 0.00012047962809447199
step: 190, loss: 0.0010632362682372332
step: 200, loss: 0.000226096817641519
step: 210, loss: 0.16101497411727905
step: 220, loss: 0.0005620406591333449
step: 230, loss: 0.0023512125480920076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932885906040269, f1=0.9833147942157954,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004637414822354913
step: 10, loss: 0.0010126791894435883
step: 20, loss: 0.0004009078838862479
step: 30, loss: 0.0003106571966782212
step: 40, loss: 0.0005889981985092163

```

```

step: 50, loss: 0.00031762002618052065
step: 60, loss: 0.0065933517180383205
step: 70, loss: 0.0004366551584098488
step: 80, loss: 0.0007407658267766237
step: 90, loss: 0.00034267205046489835
step: 100, loss: 0.006778880022466183
step: 110, loss: 0.0022345460020005703
step: 120, loss: 0.00010957615450024605
step: 130, loss: 0.0030390163883566856
step: 140, loss: 0.0006046544876880944
step: 150, loss: 0.0010794955305755138
step: 160, loss: 0.0006520047900266945
step: 170, loss: 0.010885420255362988
step: 180, loss: 0.001032702624797821
step: 190, loss: 0.0007349100196734071
step: 200, loss: 0.0005366966361179948
step: 210, loss: 0.00019460263138171285
step: 220, loss: 0.00046863191528245807
step: 230, loss: 9.501912427367643e-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 5: dev_f1=0.9887640449438202, f1=0.9854423292273236,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017544520960655063
step: 10, loss: 0.00015141525364015251
step: 20, loss: 0.0018810121109709144
step: 30, loss: 9.770248288987204e-05
step: 40, loss: 0.00034214049810543656
step: 50, loss: 0.0007508231210522354
step: 60, loss: 0.0005343233933672309
step: 70, loss: 0.00020011121523566544
step: 80, loss: 0.0002771932340692729
step: 90, loss: 0.00011045373685192317
step: 100, loss: 0.021306345239281654
step: 110, loss: 0.00014474705676548183
step: 120, loss: 0.00028843755717389286
step: 130, loss: 0.0004474532324820757
step: 140, loss: 0.03612039238214493
step: 150, loss: 0.0012943558394908905
step: 160, loss: 0.0005573760718107224
step: 170, loss: 0.0010098041966557503
step: 180, loss: 0.03051166795194149
step: 190, loss: 0.013188731856644154
step: 200, loss: 0.00017940420366358012

```

```

step: 210, loss: 0.0001607925078133121
step: 220, loss: 6.870744255138561e-05
step: 230, loss: 0.05110081285238266
/usr/local/lib/python3.7/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=0.984304932735426,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007551448070444167
step: 10, loss: 5.5902692110976204e-05
step: 20, loss: 0.0001275217509828508
step: 30, loss: 7.862936035962775e-05
step: 40, loss: 6.91083914716728e-05
step: 50, loss: 0.00015906906628515571
step: 60, loss: 0.0007162965484894812
step: 70, loss: 0.022666968405246735
step: 80, loss: 5.64155780011788e-05
step: 90, loss: 3.653207022580318e-05
step: 100, loss: 5.3215542720863596e-05
step: 110, loss: 0.00013612139446195215
step: 120, loss: 7.334105612244457e-05
step: 130, loss: 4.2008046875707805e-05
step: 140, loss: 5.447940930025652e-05
step: 150, loss: 0.007765966467559338
step: 160, loss: 0.0450914092361927
step: 170, loss: 0.0003581536584533751
step: 180, loss: 0.0005176147678866982
step: 190, loss: 0.00023339157633017749
step: 200, loss: 0.02275528758764267
step: 210, loss: 3.088885932811536e-05
step: 220, loss: 0.0003253673203289509
step: 230, loss: 0.04729078710079193
/usr/local/lib/python3.7/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=0.9910313901345291, f1=0.9844444444444443,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.6386762076290324e-05
step: 10, loss: 0.0011897593503817916
step: 20, loss: 0.04633091390132904
step: 30, loss: 0.003932317718863487
step: 40, loss: 0.09664107859134674

```



```

step: 50, loss: 0.0004003688518423587
step: 60, loss: 0.00020912520994897932
step: 70, loss: 0.00018051858933176845
step: 80, loss: 0.00012510047235991806
step: 90, loss: 0.002437157556414604
step: 100, loss: 0.0002267586241941899
step: 110, loss: 0.016718033701181412
step: 120, loss: 0.00044779846211895347
step: 130, loss: 0.00017854422912932932
step: 140, loss: 7.102829113136977e-05
step: 150, loss: 0.00010175551869906485
step: 160, loss: 0.004744641482830048
step: 170, loss: 0.00022775404795538634
step: 180, loss: 0.0014835742767900229
step: 190, loss: 0.00781903974711895
step: 200, loss: 0.00010083920642500743
step: 210, loss: 5.7742417993722484e-05
step: 220, loss: 0.00031296457746066153
step: 230, loss: 6.617651524720713e-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=0.9875706214689265, f1=0.9853438556933484,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.287838914431632e-05
step: 10, loss: 7.145877316361293e-05
step: 20, loss: 0.00030117787537164986
step: 30, loss: 6.047104034223594e-05
step: 40, loss: 0.020263848826289177
step: 50, loss: 3.415587707422674e-05
step: 60, loss: 0.00023600776330567896
step: 70, loss: 3.5664015740621835e-05
step: 80, loss: 4.377638106234372e-05
step: 90, loss: 4.869392068940215e-05
step: 100, loss: 0.00011233552504563704
step: 110, loss: 3.9303457015194e-05
step: 120, loss: 3.554917930159718e-05
step: 130, loss: 0.00010281893628416583
step: 140, loss: 0.0003432712110225111
step: 150, loss: 0.005068622529506683
step: 160, loss: 4.701867146650329e-05
step: 170, loss: 7.125687989173457e-05
step: 180, loss: 0.0001239917182829231
step: 190, loss: 7.437539898091927e-05
step: 200, loss: 8.204700861824676e-05

```

```

step: 210, loss: 7.378802547464147e-05
step: 220, loss: 4.167341467109509e-05
step: 230, loss: 6.080698949517682e-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=0.9887892376681614, f1=0.9865771812080537,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.887722752755508e-05
step: 10, loss: 5.300764678395353e-05
step: 20, loss: 5.7461300457362086e-05
step: 30, loss: 5.6276210671057925e-05
step: 40, loss: 5.218247679295018e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 4.0703333070268854e-05
step: 60, loss: 9.87654275377281e-05
step: 70, loss: 5.2115439757471904e-05
step: 80, loss: 6.557541928486899e-05
step: 90, loss: 6.306373688858002e-05
step: 100, loss: 7.041525532258675e-05
step: 110, loss: 5.531431088456884e-05
step: 120, loss: 0.00012114513083361089
step: 130, loss: 3.9072459912858903e-05
step: 140, loss: 0.03590058535337448
step: 150, loss: 0.012345312163233757
step: 160, loss: 2.5707513486850075e-05
step: 170, loss: 0.00011281643674010411
step: 180, loss: 0.00020707283692900091
step: 190, loss: 0.018023276701569557
step: 200, loss: 4.1174746002070606e-05
step: 210, loss: 2.7570000384002924e-05
step: 220, loss: 8.715520380064845e-05
step: 230, loss: 0.0001408659154549241
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9821428571428571,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.320240597939119e-05
step: 10, loss: 9.288627916248515e-05
step: 20, loss: 0.001076901564374566
step: 30, loss: 0.017512699589133263

```

```

step: 40, loss: 0.0035249374341219664
step: 50, loss: 0.00011576752149267122
step: 60, loss: 0.0002097420219797641
step: 70, loss: 0.0011199417058378458
step: 80, loss: 3.405952520552091e-05
step: 90, loss: 5.911653715884313e-05
step: 100, loss: 9.664912067819387e-05
step: 110, loss: 0.01413261890411377
step: 120, loss: 3.113143247901462e-05
step: 130, loss: 3.629372440627776e-05
step: 140, loss: 4.0062397602014244e-05
step: 150, loss: 0.03285898268222809
step: 160, loss: 3.576918606995605e-05
step: 170, loss: 0.01694546267390251
step: 180, loss: 3.5981043765787035e-05
step: 190, loss: 3.3049756893888116e-05
step: 200, loss: 0.00012196283205412328
step: 210, loss: 3.4711330954451114e-05
step: 220, loss: 2.6810415874933824e-05
step: 230, loss: 0.00010515992471482605
/usr/local/lib/python3.7/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=0.988814317673378, f1=0.9844097995545658,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.40964543283917e-05
step: 10, loss: 4.361465107649565e-05
step: 20, loss: 3.3592725230846554e-05
step: 30, loss: 3.546377411112189e-05
step: 40, loss: 9.04335465747863e-05
step: 50, loss: 4.3977503082714975e-05
step: 60, loss: 0.0001827558153308928
step: 70, loss: 3.122778070974164e-05
step: 80, loss: 5.130054705659859e-05
step: 90, loss: 3.073222251259722e-05
step: 100, loss: 3.310840475023724e-05
step: 110, loss: 3.400313653401099e-05
step: 120, loss: 3.562291749403812e-05
step: 130, loss: 3.34594733431004e-05
step: 140, loss: 0.021578893065452576
step: 150, loss: 3.4606629924383014e-05
step: 160, loss: 2.718653740885202e-05
step: 170, loss: 4.5640237658517435e-05
step: 180, loss: 0.020719241350889206
step: 190, loss: 9.927130304276943e-05

```

```

step: 200, loss: 2.1099729565321468e-05
step: 210, loss: 2.900073741329834e-05
step: 220, loss: 2.1744188416050747e-05
step: 230, loss: 0.03483269736170769
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.9855072463768116,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.008225005236454e-05
step: 10, loss: 9.415423846803606e-05
step: 20, loss: 0.0001056876135407947
step: 30, loss: 3.4509961551520973e-05
step: 40, loss: 5.922167110838927e-05
step: 50, loss: 0.0019036586163565516
step: 60, loss: 0.00012050700752297416
step: 70, loss: 3.445145193836652e-05
step: 80, loss: 4.870377961196937e-05
step: 90, loss: 0.00029088498558849096
step: 100, loss: 2.346885230508633e-05
step: 110, loss: 0.02673262171447277
step: 120, loss: 0.036823298782110214
step: 130, loss: 4.2037867387989536e-05
step: 140, loss: 3.929551530745812e-05
step: 150, loss: 3.742634362424724e-05
step: 160, loss: 0.00012656793114729226
step: 170, loss: 4.64462282252498e-05
step: 180, loss: 3.618983464548364e-05
step: 190, loss: 2.393435534031596e-05
step: 200, loss: 3.774492142838426e-05
step: 210, loss: 2.0112443962716497e-05
step: 220, loss: 2.3729613531031646e-05
step: 230, loss: 4.206542871543206e-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=0.9909706546275394, f1=0.984304932735426,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.0048798837233335e-05
step: 10, loss: 0.018580235540866852
step: 20, loss: 3.281432145740837e-05
step: 30, loss: 4.0231167076854035e-05

```

```

step: 40, loss: 0.01227030623704195
step: 50, loss: 2.9011775040999055e-05
step: 60, loss: 5.12426886416506e-05
step: 70, loss: 5.0628194003365934e-05
step: 80, loss: 2.9741790058324113e-05
step: 90, loss: 3.374580410309136e-05
step: 100, loss: 6.312446930678561e-05
step: 110, loss: 5.6806155043886974e-05
step: 120, loss: 1.6532612789887935e-05
step: 130, loss: 2.616218444018159e-05
step: 140, loss: 6.186653627082705e-05
step: 150, loss: 2.5491368432994932e-05
step: 160, loss: 0.008408892899751663
step: 170, loss: 4.0917286241892725e-05
step: 180, loss: 2.7089545255876146e-05
step: 190, loss: 5.045767466071993e-05
step: 200, loss: 2.190795203205198e-05
step: 210, loss: 3.435353210079484e-05
step: 220, loss: 8.794770110398531e-05
step: 230, loss: 2.9302260372787714e-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=0.9844097995545658,
best_f1=0.9833147942157954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.1028847186244093e-05
step: 10, loss: 1.729253381199669e-05
step: 20, loss: 2.6571900889393874e-05
step: 30, loss: 4.490624996833503e-05
step: 40, loss: 4.005571099696681e-05
step: 50, loss: 0.05029723420739174
step: 60, loss: 4.415650982991792e-05
step: 70, loss: 3.354418367962353e-05
step: 80, loss: 2.629241680551786e-05
step: 90, loss: 3.916221612598747e-05
step: 100, loss: 2.277601379319094e-05
step: 110, loss: 3.842102159978822e-05
step: 120, loss: 0.0003483067557681352
step: 130, loss: 0.018247175961732864
step: 140, loss: 2.2906422600499354e-05
step: 150, loss: 4.683176666731015e-05
step: 160, loss: 2.270869845233392e-05
step: 170, loss: 2.0216724806232378e-05
step: 180, loss: 0.0005268450477160513
step: 190, loss: 5.1386490667937323e-05

```

```

step: 200, loss: 7.489239942515269e-05
step: 210, loss: 3.421097426326014e-05
step: 220, loss: 0.00016584908007644117
step: 230, loss: 2.1792582629132085e-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=0.9844097995545658,
best_f1=0.9833147942157954
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 208.81it/s]
load_f1 = 0.9932885906040269
real_f1 = 0.9910714285714286
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 [00:17, 257.57it/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, 624kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 25.6MB/s]

Downloading: 100% 440M/440M [00:08<00:00, 52.0MB/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: 0.6126366257667542
step: 10, loss: 0.5238776803016663
step: 20, loss: 0.432565301656723
step: 30, loss: 0.04129697382450104
step: 40, loss: 0.11765676736831665
step: 50, loss: 0.040156543254852295
step: 60, loss: 0.038968365639448166
step: 70, loss: 0.07298902422189713
step: 80, loss: 0.02459871396422386
step: 90, loss: 0.14091292023658752
step: 100, loss: 0.03779961168766022
step: 110, loss: 0.09646955132484436
step: 120, loss: 0.05693099647760391
step: 130, loss: 0.09760093688964844
step: 140, loss: 0.08120217174291611
step: 150, loss: 0.034235868602991104
step: 160, loss: 0.036176566034555435
step: 170, loss: 0.15255945920944214
step: 180, loss: 0.04284001141786575
step: 190, loss: 0.00309034064412117
step: 200, loss: 0.15999294817447662
step: 210, loss: 0.09572939574718475
step: 220, loss: 0.19923186302185059

```



```

step: 230, loss: 0.15552690625190735
step: 240, loss: 0.036487143486738205
step: 250, loss: 0.016109589487314224
step: 260, loss: 0.1803824007511139
step: 270, loss: 0.00973422173410654
step: 280, loss: 0.037082698196172714
step: 290, loss: 0.0662463903427124
step: 300, loss: 0.015978123992681503
step: 310, loss: 0.21445488929748535
step: 320, loss: 0.07251245528459549
step: 330, loss: 0.004671741742640734
step: 340, loss: 0.012488136067986488
step: 350, loss: 0.04154537245631218
step: 360, loss: 0.12287680059671402
step: 370, loss: 0.07761555910110474
step: 380, loss: 0.004752027336508036
step: 390, loss: 0.23971696197986603
step: 400, loss: 0.17390412092208862
step: 410, loss: 0.05320644751191139
step: 420, loss: 0.02051590010523796
step: 430, loss: 0.14551274478435516
step: 440, loss: 0.022642815485596657
step: 450, loss: 0.007371845655143261
step: 460, loss: 0.004875205922871828
step: 470, loss: 0.1130807027220726
step: 480, loss: 0.04132857173681259
step: 490, loss: 0.06611557304859161
step: 500, loss: 0.04288548231124878
step: 510, loss: 0.051498737186193466
step: 520, loss: 0.014361868612468243
step: 530, loss: 0.003342017997056246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9414455626715461, f1=0.9396277802995915,
best_f1=0.9396277802995915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26446202397346497
step: 10, loss: 0.04114454612135887
step: 20, loss: 0.011377573944628239
step: 30, loss: 0.043938178569078445
step: 40, loss: 0.16455039381980896

```

step: 50, loss: 0.12410634011030197
step: 60, loss: 0.002890883246436715
step: 70, loss: 0.015922537073493004
step: 80, loss: 0.057451099157333374
step: 90, loss: 0.008649773895740509
step: 100, loss: 0.006654088851064444
step: 110, loss: 0.005367041565477848
step: 120, loss: 0.06489483267068863
step: 130, loss: 0.0950661301612854
step: 140, loss: 0.012695277109742165
step: 150, loss: 0.040151093155145645
step: 160, loss: 0.008846058510243893
step: 170, loss: 0.006825195159763098
step: 180, loss: 0.06164100021123886
step: 190, loss: 0.039271511137485504
step: 200, loss: 0.0064904349856078625
step: 210, loss: 0.029998701065778732
step: 220, loss: 0.08497241139411926
step: 230, loss: 0.006782948970794678
step: 240, loss: 0.1024659052491188
step: 250, loss: 0.004441050812602043
step: 260, loss: 0.001422145520336926
step: 270, loss: 0.08509036898612976
step: 280, loss: 0.007979309186339378
step: 290, loss: 0.03854750841856003
step: 300, loss: 0.13567538559436798
step: 310, loss: 0.011297102086246014
step: 320, loss: 0.06445008516311646
step: 330, loss: 0.02403312362730503
step: 340, loss: 0.0031126076355576515
step: 350, loss: 0.0010944025125354528
step: 360, loss: 0.027011679485440254
step: 370, loss: 0.10169815272092819
step: 380, loss: 0.11253886669874191
step: 390, loss: 0.04061676561832428
step: 400, loss: 0.09258078783750534
step: 410, loss: 0.046390485018491745
step: 420, loss: 0.05706591531634331
step: 430, loss: 0.0058407289907336235
step: 440, loss: 0.08183236420154572
step: 450, loss: 0.010263850912451744
step: 460, loss: 0.014956501312553883
step: 470, loss: 0.15883758664131165
step: 480, loss: 0.2459830641746521
step: 490, loss: 0.06210390850901604
step: 500, loss: 0.14699150621891022
step: 510, loss: 0.005369152873754501
step: 520, loss: 0.050356678664684296

```

step: 530, loss: 0.025507967919111252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9446254071661238, f1=0.9428172942817293,
best_f1=0.9428172942817293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0636129155755043
step: 10, loss: 0.06388124823570251
step: 20, loss: 0.03809400647878647
step: 30, loss: 0.09342887252569199
step: 40, loss: 0.013778450898826122
step: 50, loss: 0.10713383555412292
step: 60, loss: 0.008896942250430584
step: 70, loss: 0.006341141182929277
step: 80, loss: 0.002387574641034007
step: 90, loss: 0.006472235545516014
step: 100, loss: 0.09383709728717804
step: 110, loss: 0.032509222626686096
step: 120, loss: 0.005696522071957588
step: 130, loss: 0.007056492380797863
step: 140, loss: 0.06295539438724518
step: 150, loss: 0.07874217629432678
step: 160, loss: 0.0046492381952703
step: 170, loss: 0.1619957983493805
step: 180, loss: 0.08674879372119904
step: 190, loss: 0.009252380579710007
step: 200, loss: 0.002791669685393572
step: 210, loss: 0.09689949452877045
step: 220, loss: 0.028576627373695374
step: 230, loss: 0.11221467703580856
step: 240, loss: 0.0021666157990694046
step: 250, loss: 0.1350669115781784
step: 260, loss: 0.01660880446434021
step: 270, loss: 0.0016353961545974016
step: 280, loss: 0.046324945986270905
step: 290, loss: 0.0005467784940265119
step: 300, loss: 0.0060798777267336845
step: 310, loss: 0.031638696789741516
step: 320, loss: 0.08529967069625854
step: 330, loss: 0.0035012168809771538
step: 340, loss: 0.0026735193096101284

```

```

step: 350, loss: 0.0017211259109899402
step: 360, loss: 0.014988075941801071
step: 370, loss: 0.009136472828686237
step: 380, loss: 0.04738199710845947
step: 390, loss: 0.011420872993767262
step: 400, loss: 0.02879492938518524
step: 410, loss: 0.003527577966451645
step: 420, loss: 0.1407504677772522
step: 430, loss: 0.01187664270401001
step: 440, loss: 0.009640648029744625
step: 450, loss: 0.0861581563949585
step: 460, loss: 0.10798293352127075
step: 470, loss: 0.035126976668834686
step: 480, loss: 0.03744044899940491
step: 490, loss: 0.003124591428786516
step: 500, loss: 0.040640827268362045
step: 510, loss: 0.003831257112324238
step: 520, loss: 0.05142422020435333
step: 530, loss: 0.013392709195613861
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9467401285583104, f1=0.9515096065873743,
best_f1=0.9515096065873743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005284568294882774
step: 10, loss: 0.009995927102863789
step: 20, loss: 0.0008477535448037088
step: 30, loss: 0.0008506090380251408
step: 40, loss: 0.0007969877915456891
step: 50, loss: 0.002971082227304578
step: 60, loss: 0.0007280526333488524
step: 70, loss: 0.02774285525083542
step: 80, loss: 0.03849007934331894
step: 90, loss: 0.02639615163207054
step: 100, loss: 0.008049539290368557
step: 110, loss: 0.011613699607551098
step: 120, loss: 0.00013507074618246406
step: 130, loss: 0.0021783336997032166
step: 140, loss: 0.0009481866145506501
step: 150, loss: 0.09599748253822327
step: 160, loss: 0.19663512706756592

```

```

step: 170, loss: 0.007428137119859457
step: 180, loss: 0.0017281591426581144
step: 190, loss: 0.14486472308635712
step: 200, loss: 0.0031390299554914236
step: 210, loss: 0.044503554701805115
step: 220, loss: 0.013123405165970325
step: 230, loss: 0.052611093968153
step: 240, loss: 0.007136571686714888
step: 250, loss: 0.011176844127476215
step: 260, loss: 0.000545137096196413
step: 270, loss: 0.009691640734672546
step: 280, loss: 0.08577536046504974
step: 290, loss: 0.014960154891014099
step: 300, loss: 0.0001945091935340315
step: 310, loss: 0.01951863057911396
step: 320, loss: 0.00774215767160058
step: 330, loss: 0.030542420223355293
step: 340, loss: 0.005786184687167406
step: 350, loss: 0.00997857004404068
step: 360, loss: 0.02361958473920822
step: 370, loss: 0.008634347468614578
step: 380, loss: 0.08050628751516342
step: 390, loss: 0.004014445934444666
step: 400, loss: 0.031064148992300034
step: 410, loss: 0.0019547350239008665
step: 420, loss: 0.0015798215754330158
step: 430, loss: 0.04938637092709541
step: 440, loss: 0.00044938427163287997
step: 450, loss: 0.012484809383749962
step: 460, loss: 0.002876771381124854
step: 470, loss: 0.16928640007972717
step: 480, loss: 0.022182155400514603
step: 490, loss: 0.0011398355709388852
step: 500, loss: 0.0031967225950211287
step: 510, loss: 0.0043510752730071545
step: 520, loss: 0.1353301852941513
step: 530, loss: 0.02260514348745346
/usr/local/lib/python3.7/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=0.937095282146161, f1=0.9462563160312356,
best_f1=0.9515096065873743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04156701639294624
step: 10, loss: 0.061955973505973816
step: 20, loss: 0.005016889423131943

```

step: 30, loss: 0.0007413060520775616
step: 40, loss: 0.10200564563274384
step: 50, loss: 0.009374119341373444
step: 60, loss: 0.06342402845621109
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 70, loss: 0.006322719156742096
step: 80, loss: 0.0004550896701402962
step: 90, loss: 0.0009574479190632701
step: 100, loss: 0.025476643815636635
step: 110, loss: 0.0003841875004582107
step: 120, loss: 0.0023077442310750484
step: 130, loss: 0.000296242069453001
step: 140, loss: 0.0005330815329216421
step: 150, loss: 0.00564236706122756
step: 160, loss: 0.0002899573009926826
step: 170, loss: 0.035191237926483154
step: 180, loss: 0.0013938096817582846
step: 190, loss: 0.0005913585773669183
step: 200, loss: 0.0005412191967479885
step: 210, loss: 0.001641340903006494
step: 220, loss: 0.0007430568221025169
step: 230, loss: 0.0023350841365754604
step: 240, loss: 0.0033990193624049425
step: 250, loss: 0.0021571991965174675
step: 260, loss: 0.0015190759440883994
step: 270, loss: 7.193559576990083e-05
step: 280, loss: 0.00014407529670279473
step: 290, loss: 0.14108525216579437
step: 300, loss: 0.0029955157078802586
step: 310, loss: 0.0046656630001962185
step: 320, loss: 0.19128121435642242
step: 330, loss: 0.006023306865245104
step: 340, loss: 0.00949384830892086
step: 350, loss: 0.003459192579612136
step: 360, loss: 0.0011215852573513985
step: 370, loss: 0.0023861562367528677
step: 380, loss: 0.00012311500904615968
step: 390, loss: 0.00035584327997639775
step: 400, loss: 0.001233168295584619
step: 410, loss: 0.0606473833322525
step: 420, loss: 0.0006227233679965138
step: 430, loss: 0.0009772221092134714
step: 440, loss: 0.08493304252624512
step: 450, loss: 0.002911314368247986
step: 460, loss: 8.323093788931146e-05
step: 470, loss: 0.014324271120131016
step: 480, loss: 0.0015128368977457285
step: 490, loss: 0.00015677472401876003

```

step: 500, loss: 0.00075701525202021
step: 510, loss: 0.11123456060886383
step: 520, loss: 0.000281979504507035
step: 530, loss: 0.0016095441533252597
/usr/local/lib/python3.7/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=0.9431870669745959, f1=0.9440879926672778,
best_f1=0.9515096065873743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04215503856539726
step: 10, loss: 0.0006979461177252233
step: 20, loss: 0.09857107698917389
step: 30, loss: 0.0007335032569244504
step: 40, loss: 0.011540358886122704
step: 50, loss: 0.003967329394072294
step: 60, loss: 0.00029822613578289747
step: 70, loss: 0.000300941668683663
step: 80, loss: 0.008038273081183434
step: 90, loss: 0.0005747203249484301
step: 100, loss: 0.003071252955123782
step: 110, loss: 0.0007802189793437719
step: 120, loss: 0.00040195343899540603
step: 130, loss: 0.00263035437092185
step: 140, loss: 0.0021897684782743454
step: 150, loss: 0.0022973378654569387
step: 160, loss: 0.0012031800579279661
step: 170, loss: 0.0011126056779175997
step: 180, loss: 0.0013960959622636437
step: 190, loss: 0.0003074277192354202
step: 200, loss: 4.778106085723266e-05
step: 210, loss: 0.002248705830425024
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.0009198529878631234
step: 230, loss: 0.0008575060637667775
step: 240, loss: 0.002506396034732461
step: 250, loss: 0.0010288696503266692
step: 260, loss: 8.145230094669387e-05
step: 270, loss: 0.05646631866693497
step: 280, loss: 0.0018615564331412315
step: 290, loss: 0.00027404428692534566
step: 300, loss: 0.028240254148840904
step: 310, loss: 0.0018463422311469913
step: 320, loss: 0.022967511788010597
step: 330, loss: 0.0006106351502239704
step: 340, loss: 0.059173379093408585

```

```

step: 350, loss: 0.002250713761895895
step: 360, loss: 0.0093708336353302
step: 370, loss: 0.0018208589171990752
step: 380, loss: 0.001467209542170167
step: 390, loss: 0.012453563511371613
step: 400, loss: 0.00187861907761544
step: 410, loss: 0.001402412774041295
step: 420, loss: 0.059432923793792725
step: 430, loss: 0.004148220643401146
step: 440, loss: 0.0010570805752649903
step: 450, loss: 0.0015367217129096389
step: 460, loss: 6.245973781915382e-05
step: 470, loss: 0.0006273635663092136
step: 480, loss: 0.01853281632065773
step: 490, loss: 0.009136538952589035
step: 500, loss: 0.012642890214920044
step: 510, loss: 0.003379515837877989
step: 520, loss: 0.0016708095790818334
step: 530, loss: 0.007127393968403339
/usr/local/lib/python3.7/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=0.9394650398873767, f1=0.9473684210526316,
best_f1=0.9515096065873743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002492188010364771
step: 10, loss: 0.0008893431513570249
step: 20, loss: 0.0016953549347817898
step: 30, loss: 0.0012204934610053897
step: 40, loss: 0.0010939326602965593
step: 50, loss: 0.010832234285771847
step: 60, loss: 0.04870923236012459
step: 70, loss: 0.0019998657517135143
step: 80, loss: 0.024293262511491776
step: 90, loss: 0.0007183091947808862
step: 100, loss: 0.00010254168591927737
step: 110, loss: 0.0006570973200723529
step: 120, loss: 0.0016858682502061129
step: 130, loss: 0.010648640803992748
step: 140, loss: 4.505673859966919e-05
step: 150, loss: 0.00032129729515872896
step: 160, loss: 7.137977809179574e-05
step: 170, loss: 5.182747918297537e-05
step: 180, loss: 0.015338330529630184
step: 190, loss: 0.0028810605872422457
step: 200, loss: 0.0033085744362324476

```



```

step: 210, loss: 0.005039269104599953
step: 220, loss: 4.570386226987466e-05
step: 230, loss: 0.04207904264330864
step: 240, loss: 0.01599700003862381
step: 250, loss: 0.007856767624616623
step: 260, loss: 0.07162673771381378
step: 270, loss: 4.950480433763005e-05
step: 280, loss: 0.12244054675102234
step: 290, loss: 0.00039895527879707515
step: 300, loss: 0.12414796650409698
step: 310, loss: 0.0018071138765662909
step: 320, loss: 0.00017585283785592765
step: 330, loss: 0.020378364250063896
step: 340, loss: 0.022988315671682358
step: 350, loss: 0.0004218388057779521
step: 360, loss: 0.0077401986345648766
step: 370, loss: 0.00011182448361068964
step: 380, loss: 0.005208652000874281
step: 390, loss: 0.0005320406053215265
step: 400, loss: 0.014568147249519825
step: 410, loss: 0.0019882931374013424
step: 420, loss: 0.00016856052388902754
step: 430, loss: 0.00047088865539990366
step: 440, loss: 0.0018355357460677624
step: 450, loss: 0.0005013567279092968
step: 460, loss: 0.0011553397634997964
step: 470, loss: 0.003875207155942917
step: 480, loss: 0.03019988164305687
step: 490, loss: 0.00024602757184766233
step: 500, loss: 0.004029380157589912
step: 510, loss: 0.001449773320928216
step: 520, loss: 0.027365684509277344
step: 530, loss: 0.0001151167307398282
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9484918793503482, f1=0.9456221198156683,
best_f1=0.9456221198156683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003484518965706229
step: 10, loss: 0.016242384910583496
step: 20, loss: 0.00031329336343333125

```

step: 30, loss: 0.0007409653626382351
step: 40, loss: 3.944090349250473e-05
step: 50, loss: 0.0005330797866918147
step: 60, loss: 0.0024980490561574697
step: 70, loss: 0.00010021651542047039
step: 80, loss: 0.0001074976462405175
step: 90, loss: 0.0006266233394853771
step: 100, loss: 2.7581190806813538e-05
step: 110, loss: 0.00026586753665469587
step: 120, loss: 8.652519318275154e-05
step: 130, loss: 0.0003800051927100867
step: 140, loss: 0.00803989265114069
step: 150, loss: 0.00032917928183451295
step: 160, loss: 0.06516993045806885
step: 170, loss: 0.0013786789495497942
step: 180, loss: 2.4206499801948667e-05
step: 190, loss: 0.0018249668646603823
step: 200, loss: 0.0008948339382186532
step: 210, loss: 0.0001322804018855095
step: 220, loss: 0.0016076673055067658
step: 230, loss: 2.0891069652861916e-05
step: 240, loss: 6.022996240062639e-05
step: 250, loss: 2.5923647626768798e-05
step: 260, loss: 0.00015958643052726984
step: 270, loss: 0.00011770728451665491
step: 280, loss: 0.00891882088035345
step: 290, loss: 0.00012653300655074418
step: 300, loss: 0.0006520124152302742
step: 310, loss: 8.467580482829362e-05
step: 320, loss: 0.10479484498500824
step: 330, loss: 0.0004991164896637201
step: 340, loss: 3.95374845538754e-05
step: 350, loss: 0.06848128139972687
step: 360, loss: 0.0035998427774757147
step: 370, loss: 3.775686491280794e-05
step: 380, loss: 0.00985698401927948
step: 390, loss: 0.14696429669857025
step: 400, loss: 3.254256080253981e-05
step: 410, loss: 0.0011544468579813838
step: 420, loss: 0.0038283541798591614
step: 430, loss: 0.05858493968844414
step: 440, loss: 0.002515649190172553
step: 450, loss: 0.012091600336134434
step: 460, loss: 0.043273620307445526
step: 470, loss: 0.00023488195438403636
step: 480, loss: 4.380774771561846e-05
step: 490, loss: 0.0002456710208207369
step: 500, loss: 0.00040429699583910406

```

step: 510, loss: 0.00039111581281758845
step: 520, loss: 0.0014609801582992077
step: 530, loss: 0.0017342722276225686
/usr/local/lib/python3.7/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=0.9383624655013799, f1=0.9371847776249428,
best_f1=0.9456221198156683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029815970920026302
step: 10, loss: 0.0006826804019510746
step: 20, loss: 0.008174263872206211
step: 30, loss: 0.0016516498290002346
step: 40, loss: 0.0007667164318263531
step: 50, loss: 2.5089102564379573e-05
step: 60, loss: 0.004084552638232708
step: 70, loss: 0.0016020929906517267
step: 80, loss: 0.004321770742535591
step: 90, loss: 2.674690040294081e-05
step: 100, loss: 0.0004389073292259127
step: 110, loss: 2.229912752227392e-05
step: 120, loss: 0.002987053245306015
step: 130, loss: 0.013841813430190086
step: 140, loss: 3.210249633411877e-05
step: 150, loss: 2.1710689907195047e-05
step: 160, loss: 0.0014966827584430575
step: 170, loss: 0.014055106788873672
step: 180, loss: 7.055362948449329e-05
step: 190, loss: 2.4366601792280562e-05
step: 200, loss: 6.990852853050455e-05
step: 210, loss: 3.258843571529724e-05
step: 220, loss: 0.00018224684754386544
step: 230, loss: 0.0027020208071917295
step: 240, loss: 2.4783901608316228e-05
step: 250, loss: 8.320448250742629e-05
step: 260, loss: 0.013272231444716454
step: 270, loss: 1.9147793864249252e-05
step: 280, loss: 0.00014275286230258644
step: 290, loss: 0.02164890058338642
step: 300, loss: 2.6251535018673167e-05
step: 310, loss: 0.013739733025431633
step: 320, loss: 0.00237606861628592
step: 330, loss: 4.4527951104100794e-05
step: 340, loss: 3.660824950202368e-05
step: 350, loss: 0.0014817798510193825
step: 360, loss: 2.3077422156347893e-05

```

```

step: 370, loss: 0.00029449109570123255
step: 380, loss: 2.7391317416913807e-05
step: 390, loss: 1.8454857126926072e-05
step: 400, loss: 0.002578780287876725
step: 410, loss: 0.00019508569675963372
step: 420, loss: 0.0001648373290663585
step: 430, loss: 0.00010459836630616337
step: 440, loss: 0.0035153671633452177
step: 450, loss: 7.407281373161823e-05
step: 460, loss: 0.00013336898700799793
step: 470, loss: 1.9315306417411193e-05
step: 480, loss: 0.0005443977424874902
step: 490, loss: 0.020871561020612717
step: 500, loss: 0.0032423362135887146
step: 510, loss: 0.00024224458320531994
step: 520, loss: 0.00031265252619050443
step: 530, loss: 9.382205462316051e-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=0.9455719557195572, f1=0.9456221198156683,
best_f1=0.9456221198156683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.676612686831504e-05
step: 10, loss: 4.16482835134957e-05
step: 20, loss: 1.6923797375056893e-05
step: 30, loss: 1.9762312149396166e-05
step: 40, loss: 0.00025402498431503773
step: 50, loss: 0.0013350638328120112
step: 60, loss: 5.2015766414115205e-05
step: 70, loss: 2.566241346357856e-05
step: 80, loss: 1.7020682207657956e-05
step: 90, loss: 0.0024304469116032124
step: 100, loss: 3.03633223666111e-05
step: 110, loss: 1.6346419215551578e-05
step: 120, loss: 2.6597477699397132e-05
step: 130, loss: 1.8913076928583905e-05
step: 140, loss: 0.011141630820930004
step: 150, loss: 6.944268534425646e-05
step: 160, loss: 3.074534106417559e-05
step: 170, loss: 0.0002501103444956243
step: 180, loss: 2.0831499568885192e-05
step: 190, loss: 9.734838386066258e-05
step: 200, loss: 0.0005798038328066468
step: 210, loss: 1.9304174202261493e-05
step: 220, loss: 0.20398660004138947

```

```

step: 230, loss: 0.00011698775051627308
step: 240, loss: 0.00033003053977154195
step: 250, loss: 3.2389409170718864e-05
step: 260, loss: 0.001410720287822187
step: 270, loss: 0.0015736679779365659
step: 280, loss: 5.459694875753485e-05
step: 290, loss: 3.321623444207944e-05
step: 300, loss: 6.605053204111755e-05
step: 310, loss: 4.240028647473082e-05
step: 320, loss: 0.0016496293246746063
step: 330, loss: 7.806458597769961e-05
step: 340, loss: 5.030931060900912e-05
step: 350, loss: 4.294455357012339e-05
step: 360, loss: 7.775911944918334e-05
step: 370, loss: 0.003754890291020274
step: 380, loss: 0.008376696147024632
step: 390, loss: 0.001696932828053832
step: 400, loss: 0.0017455712659284472
step: 410, loss: 0.002403771737590432
step: 420, loss: 9.198926272802055e-05
step: 430, loss: 3.2778025342850015e-05
step: 440, loss: 0.0010619009844958782
step: 450, loss: 0.0003403390001039952
step: 460, loss: 4.325184636400081e-05
step: 470, loss: 0.00026328134117648005
step: 480, loss: 2.023920023930259e-05
step: 490, loss: 4.0208702557720244e-05
step: 500, loss: 0.0016988386632874608
step: 510, loss: 0.0012258280767127872
step: 520, loss: 8.602791785961017e-05
step: 530, loss: 0.0012835606466978788
/usr/local/lib/python3.7/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=0.9437675726335519, f1=0.9441860465116279,
best_f1=0.9456221198156683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008930638432502747
step: 10, loss: 5.1224560593254864e-05
step: 20, loss: 4.765653648064472e-05
step: 30, loss: 3.900900628650561e-05
step: 40, loss: 3.072068648179993e-05
step: 50, loss: 0.0001359043235424906
step: 60, loss: 0.026971645653247833
step: 70, loss: 2.204204247391317e-05
step: 80, loss: 0.000758458161726594

```

```
step: 90, loss: 1.8562823242973536e-05
step: 100, loss: 0.0019013317069038749
step: 110, loss: 0.00022651316248811781
step: 120, loss: 3.116732113994658e-05
step: 130, loss: 1.54560857481556e-05
step: 140, loss: 0.00719604454934597
step: 150, loss: 0.00019382996833883226
step: 160, loss: 2.0525778381852433e-05
step: 170, loss: 0.002575413789600134
step: 180, loss: 5.6982105888891965e-05
step: 190, loss: 0.001453501172363758
step: 200, loss: 0.007698704954236746
step: 210, loss: 6.535789725603536e-05
step: 220, loss: 0.0019507824908941984
step: 230, loss: 2.8950995329068974e-05
step: 240, loss: 0.0006675602635368705
step: 250, loss: 1.9344710381119512e-05
step: 260, loss: 8.888912270776927e-05
step: 270, loss: 6.314245547400787e-05
step: 280, loss: 0.00017620775906834751
step: 290, loss: 3.2232881494564936e-05
step: 300, loss: 0.0002303508808836341
step: 310, loss: 0.00015259700012393296
step: 320, loss: 0.00016851107648108155
step: 330, loss: 2.5189750886056572e-05
step: 340, loss: 0.00024824723368510604
step: 350, loss: 2.962076359835919e-05
step: 360, loss: 1.4863771866657771e-05
step: 370, loss: 0.0011041322723031044
step: 380, loss: 1.8067288692691363e-05
step: 390, loss: 1.6487916582264006e-05
step: 400, loss: 1.625692493689712e-05
step: 410, loss: 2.064135333057493e-05
step: 420, loss: 6.801367999287322e-05
step: 430, loss: 1.5254916434059851e-05
step: 440, loss: 0.0001759451552061364
step: 450, loss: 0.0003437558189034462
step: 460, loss: 7.673553773202002e-05
step: 470, loss: 3.115921936114319e-05
step: 480, loss: 0.0011407864512875676
step: 490, loss: 0.0007957842899486423
step: 500, loss: 3.225894397473894e-05
step: 510, loss: 6.755616050213575e-05
step: 520, loss: 2.01644106709864e-05
step: 530, loss: 1.4532163731928449e-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=0.9492314857941314, f1=0.943239501615136,
best_f1=0.943239501615136
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.753091419232078e-05
step: 10, loss: 1.5288416761904955e-05
step: 20, loss: 0.00010843161726370454
step: 30, loss: 0.004417176358401775
step: 40, loss: 0.00012005162716377527
step: 50, loss: 0.044850774109363556
step: 60, loss: 0.0012504446785897017
step: 70, loss: 4.298646672395989e-05
step: 80, loss: 2.201540337409824e-05
step: 90, loss: 0.00026717816945165396
step: 100, loss: 1.924808202602435e-05
step: 110, loss: 3.4864453482441604e-05
step: 120, loss: 1.438319213775685e-05
step: 130, loss: 0.0009996418375521898
step: 140, loss: 4.2137664422625676e-05
step: 150, loss: 1.3101748663757462e-05
step: 160, loss: 1.252061156264972e-05
step: 170, loss: 0.00011300882761133835
step: 180, loss: 4.6028519136598334e-05
step: 190, loss: 0.00016632629558444023
step: 200, loss: 1.573914232722018e-05
step: 210, loss: 1.4975513295212295e-05
step: 220, loss: 1.902846634038724e-05
step: 230, loss: 1.3042136743024457e-05
step: 240, loss: 2.6408508347230963e-05
step: 250, loss: 1.992215220525395e-05
step: 260, loss: 1.694973798294086e-05
step: 270, loss: 1.716220867820084e-05
step: 280, loss: 0.0006622890359722078
step: 290, loss: 6.387501343851909e-05
step: 300, loss: 0.0019756434485316277
step: 310, loss: 2.333013435418252e-05
step: 320, loss: 1.211456492455909e-05
step: 330, loss: 1.3764813957095612e-05
step: 340, loss: 0.00048239281750284135
step: 350, loss: 2.5635417841840535e-05
step: 360, loss: 0.0009052249952219427
step: 370, loss: 0.00014541353448294103
step: 380, loss: 2.146396400348749e-05

```

```

step: 390, loss: 0.0005681233596988022
step: 400, loss: 1.8905233446275815e-05
step: 410, loss: 0.002578165614977479
step: 420, loss: 0.002695259638130665
step: 430, loss: 0.013191014528274536
step: 440, loss: 1.4062843547435477e-05
step: 450, loss: 2.085352025460452e-05
step: 460, loss: 1.220769263454713e-05
step: 470, loss: 1.377600619889563e-05
step: 480, loss: 1.579505442350637e-05
step: 490, loss: 3.586632738006301e-05
step: 500, loss: 1.4107525203144178e-05
step: 510, loss: 0.0009332027402706444
step: 520, loss: 1.3343887076189276e-05
step: 530, loss: 1.4800441022089217e-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=0.93989588263133, f1=0.9433255269320844,
best_f1=0.943239501615136
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012367696035653353
step: 10, loss: 1.2587671335495543e-05
step: 20, loss: 1.1052869012928568e-05
step: 30, loss: 2.997005867655389e-05
step: 40, loss: 1.3716400644625537e-05
step: 50, loss: 3.306286453153007e-05
step: 60, loss: 1.8178945538238622e-05
step: 70, loss: 1.6830301319714636e-05
step: 80, loss: 8.893999620340765e-05
step: 90, loss: 0.034314706921577454
step: 100, loss: 0.00010168665903620422
step: 110, loss: 1.7586693502380513e-05
step: 120, loss: 1.517296095698839e-05
step: 130, loss: 1.3850513823854271e-05
step: 140, loss: 1.2677076483669225e-05
step: 150, loss: 2.1028628907515667e-05
step: 160, loss: 1.2643544323509559e-05
step: 170, loss: 0.00010201022087130696
step: 180, loss: 3.558631578925997e-05
step: 190, loss: 1.4301273040473461e-05
step: 200, loss: 1.38765881274594e-05
step: 210, loss: 1.2483359569159802e-05
step: 220, loss: 1.1827707567135803e-05
step: 230, loss: 1.1470101526356302e-05
step: 240, loss: 1.5266061382135376e-05

```



```

step: 250, loss: 1.1939475371036679e-05
step: 260, loss: 1.224866082338849e-05
step: 270, loss: 1.087778673536377e-05
step: 280, loss: 1.2349248208920471e-05
step: 290, loss: 5.079484617454e-05
step: 300, loss: 0.0005403630784712732
step: 310, loss: 0.0005866721621714532
step: 320, loss: 0.002765809651464224
step: 330, loss: 0.001328739570453763
step: 340, loss: 1.3220928849477787e-05
step: 350, loss: 3.258126525906846e-05
step: 360, loss: 1.2751566828228533e-05
step: 370, loss: 1.5679370335419662e-05
step: 380, loss: 1.4256538634072058e-05
step: 390, loss: 0.0012519739102572203
step: 400, loss: 1.317244823439978e-05
step: 410, loss: 0.0018156617879867554
step: 420, loss: 2.588321513030678e-05
step: 430, loss: 0.00010645773727446795
step: 440, loss: 2.33748578466475e-05
step: 450, loss: 2.7447980755823664e-05
step: 460, loss: 0.0008817928028292954
step: 470, loss: 1.1660008567559998e-05
step: 480, loss: 2.3274144041351974e-05
step: 490, loss: 5.0775535783031955e-05
step: 500, loss: 2.754859087872319e-05
step: 510, loss: 9.376807429362088e-05
step: 520, loss: 0.00029195004026405513
step: 530, loss: 1.3179904271964915e-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=0.9424964936886395, f1=0.9436749769159741,
best_f1=0.943239501615136
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010183510603383183
step: 10, loss: 9.890591172734275e-06
step: 20, loss: 1.7832036974141374e-05
step: 30, loss: 1.257272651855601e-05
step: 40, loss: 2.272656638524495e-05
step: 50, loss: 0.003478679805994034
step: 60, loss: 9.36905689741252e-06
step: 70, loss: 8.903404523152858e-05
step: 80, loss: 0.003296838141977787
step: 90, loss: 1.4755500160390511e-05
step: 100, loss: 1.7981394194066525e-05

```

```

step: 110, loss: 1.0959734936477616e-05
step: 120, loss: 1.1514783182065003e-05
step: 130, loss: 0.0422229990363121
step: 140, loss: 0.001277297385968268
step: 150, loss: 9.115741704590619e-06
step: 160, loss: 0.0007899744668975472
step: 170, loss: 2.1835370716871694e-05
step: 180, loss: 1.2352954399830196e-05
step: 190, loss: 1.57349986693589e-05
step: 200, loss: 1.41335967782652e-05
step: 210, loss: 1.0766018021968193e-05
step: 220, loss: 1.1045418432331644e-05
step: 230, loss: 1.5843401342863217e-05
step: 240, loss: 1.4613895473303273e-05
step: 250, loss: 1.1429113328631502e-05
step: 260, loss: 0.0010989997535943985
step: 270, loss: 1.1205595910723787e-05
step: 280, loss: 1.0602114343782887e-05
step: 290, loss: 0.0008518323302268982
step: 300, loss: 8.814137254375964e-05
step: 310, loss: 8.322645589942113e-05
step: 320, loss: 0.0008951954077929258
step: 330, loss: 4.2132287489948794e-05
step: 340, loss: 2.7124515327159315e-05
step: 350, loss: 2.2357011403073557e-05
step: 360, loss: 0.0011658509029075503
step: 370, loss: 9.823537766351365e-06
step: 380, loss: 1.251313278771704e-05
step: 390, loss: 0.05097588896751404
step: 400, loss: 6.411375943571329e-05
step: 410, loss: 9.320627214037813e-06
step: 420, loss: 1.1201856978004798e-05
step: 430, loss: 3.931925311917439e-05
step: 440, loss: 4.086640547029674e-05
step: 450, loss: 1.8193859432358295e-05
step: 460, loss: 0.0014521947596222162
step: 470, loss: 9.68942458712263e-06
step: 480, loss: 0.003514142706990242
step: 490, loss: 1.0467995707585942e-05
step: 500, loss: 1.8044403987005353e-05
step: 510, loss: 1.2587614037329331e-05
step: 520, loss: 0.0014160522259771824
step: 530, loss: 0.0644398033618927
/usr/local/lib/python3.7/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=0.9412317818523741, f1=0.9422180801491147,
best_f1=0.943239501615136

```

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

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

```
step: 0, loss: 1.0039594599220436e-05  
step: 10, loss: 1.0110376024385914e-05  
step: 20, loss: 1.3440605471259914e-05  
step: 30, loss: 0.00015616952441632748  
step: 40, loss: 0.08426929265260696  
step: 50, loss: 1.5791163605172187e-05  
step: 60, loss: 8.545776836399455e-06  
step: 70, loss: 0.0008856369531713426  
step: 80, loss: 1.0628178642946295e-05  
step: 90, loss: 9.484537258686032e-06  
step: 100, loss: 1.15445354822441e-05  
step: 110, loss: 0.0013099855277687311  
step: 120, loss: 1.3902504178986419e-05  
step: 130, loss: 0.09049838781356812  
step: 140, loss: 2.02571081899805e-05  
step: 150, loss: 1.3001133083889727e-05  
step: 160, loss: 1.1928259482374415e-05  
step: 170, loss: 1.6692261851858348e-05  
step: 180, loss: 8.910849828680512e-06  
step: 190, loss: 1.1056588846258819e-05  
step: 200, loss: 8.925751899369061e-06  
step: 210, loss: 0.0014573953812941909  
step: 220, loss: 8.922027518565301e-06  
step: 230, loss: 0.00016432502889074385  
step: 240, loss: 2.7570042220759206e-05  
step: 250, loss: 1.2129411516070832e-05  
step: 260, loss: 3.0102371965767816e-05  
step: 270, loss: 1.204373620566912e-05  
step: 280, loss: 1.0263112926622853e-05  
step: 290, loss: 8.594204700784758e-06  
step: 300, loss: 9.845877684711013e-06  
step: 310, loss: 1.5288243957911618e-05  
step: 320, loss: 0.00040212884778156877  
step: 330, loss: 1.9932265786337666e-05  
step: 340, loss: 1.023329150484642e-05  
step: 350, loss: 1.6081068679341115e-05  
step: 360, loss: 0.0017082318663597107  
step: 370, loss: 8.948104550654534e-06  
step: 380, loss: 0.0022148636635392904  
step: 390, loss: 1.0982073945342563e-05  
step: 400, loss: 1.0132721399713773e-05  
step: 410, loss: 1.461754072806798e-05  
step: 420, loss: 1.4327190001495183e-05  
step: 430, loss: 0.0005739623447880149  
step: 440, loss: 3.767587986658327e-05
```

```

step: 450, loss: 8.325988346769009e-06
step: 460, loss: 9.011433576233685e-06
step: 470, loss: 0.0016343911411240697
step: 480, loss: 8.709686881047674e-06
step: 490, loss: 8.828896170598455e-06
step: 500, loss: 0.0003714091726578772
step: 510, loss: 1.1771776371460874e-05
step: 520, loss: 9.536693141853902e-06
step: 530, loss: 0.0012759894598275423
/usr/local/lib/python3.7/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=0.9388915206063476, f1=0.9384687646782527,
best_f1=0.943239501615136
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, 260.46it/s]
load_f1 = 0.9437470943747095
real_f1 = 0.9427640763145649
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 [00:17, 258.14it/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: 0.5526285767555237
step: 10, loss: 0.3658443093299866
step: 20, loss: 0.3898921608924866
step: 30, loss: 0.33500394225120544
step: 40, loss: 0.17980459332466125
step: 50, loss: 0.38701775670051575
step: 60, loss: 0.21816450357437134
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.16987979412078857
step: 80, loss: 0.21670933067798615
step: 90, loss: 0.3075849115848541
step: 100, loss: 0.36805179715156555
step: 110, loss: 0.23593012988567352
step: 120, loss: 0.2030085027217865
step: 130, loss: 0.18062633275985718
step: 140, loss: 0.19505879282951355
step: 150, loss: 0.21974052488803864
step: 160, loss: 0.3177012503147125
step: 170, loss: 0.2796823978424072
step: 180, loss: 0.09984277933835983
step: 190, loss: 0.2384423464536667

```

```

step: 200, loss: 0.1828683316707611
step: 210, loss: 0.17503640055656433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.64453125, f1=0.6707317073170733, best_f1=0.6707317073170733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09834868460893631
step: 10, loss: 0.2302396297454834
step: 20, loss: 0.19459077715873718
step: 30, loss: 0.2948843836784363
step: 40, loss: 0.2301836907863617
step: 50, loss: 0.1076512560248375
step: 60, loss: 0.33895671367645264
step: 70, loss: 0.1726413518190384
step: 80, loss: 0.15190725028514862
step: 90, loss: 0.12424477189779282
step: 100, loss: 0.016679830849170685
step: 110, loss: 0.0855301097035408
step: 120, loss: 0.14371030032634735
step: 130, loss: 0.04855203256011009
step: 140, loss: 0.2416873723268509
step: 150, loss: 0.2824406027793884
step: 160, loss: 0.2615358233451843
step: 170, loss: 0.08499232679605484
step: 180, loss: 0.15493589639663696
step: 190, loss: 0.15596936643123627
step: 200, loss: 0.05405578017234802
step: 210, loss: 0.1207612007856369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6690140845070424, f1=0.7035714285714285,
best_f1=0.7035714285714285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04272080585360527

```

```

step: 10, loss: 0.1803630292415619
step: 20, loss: 0.06797708570957184
step: 30, loss: 0.12174967676401138
step: 40, loss: 0.19925056397914886
step: 50, loss: 0.051725469529628754
step: 60, loss: 0.1097981333732605
step: 70, loss: 0.03482639044523239
step: 80, loss: 0.1587415337562561
step: 90, loss: 0.07984161376953125
step: 100, loss: 0.1327521651983261
step: 110, loss: 0.1876770257949829
step: 120, loss: 0.24258211255073547
step: 130, loss: 0.17190226912498474
step: 140, loss: 0.15338043868541718
step: 150, loss: 0.2148548811674118
step: 160, loss: 0.02912062034010887
step: 170, loss: 0.0894210934638977
step: 180, loss: 0.11060728877782822
step: 190, loss: 0.2186940759420395
step: 200, loss: 0.029146282002329826
step: 210, loss: 0.16183757781982422
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.6915520628683693, f1=0.7061143984220907,
best_f1=0.7061143984220907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1916486620903015
step: 10, loss: 0.07330488413572311
step: 20, loss: 0.05611150339245796
step: 30, loss: 0.1437196135520935
step: 40, loss: 0.034600622951984406
step: 50, loss: 0.09925514459609985
step: 60, loss: 0.1553969532251358
step: 70, loss: 0.17686648666858673
step: 80, loss: 0.042028576135635376
step: 90, loss: 0.08495700359344482
step: 100, loss: 0.2293749451637268
step: 110, loss: 0.17388689517974854
step: 120, loss: 0.12041831761598587
step: 130, loss: 0.22707131505012512
step: 140, loss: 0.17318245768547058

```



```

step: 150, loss: 0.019804174080491066
step: 160, loss: 0.06733452528715134
step: 170, loss: 0.09833673387765884
step: 180, loss: 0.30731672048568726
step: 190, loss: 0.07163979113101959
step: 200, loss: 0.32024553418159485
step: 210, loss: 0.1471298784017563
/usr/local/lib/python3.7/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=0.6780952380952381, f1=0.7100371747211897,
best_f1=0.7061143984220907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1182909831404686
step: 10, loss: 0.02047043666243553
step: 20, loss: 0.22813951969146729
step: 30, loss: 0.1029927209019661
step: 40, loss: 0.13775520026683807
step: 50, loss: 0.10114838927984238
step: 60, loss: 0.12614531815052032
step: 70, loss: 0.08380024880170822
step: 80, loss: 0.05634523183107376
step: 90, loss: 0.12504003942012787
step: 100, loss: 0.02135474421083927
step: 110, loss: 0.11190255731344223
step: 120, loss: 0.1586594581604004
step: 130, loss: 0.034140825271606445
step: 140, loss: 0.1106186956167221
step: 150, loss: 0.08288134634494781
step: 160, loss: 0.17619448900222778
step: 170, loss: 0.04603132605552673
step: 180, loss: 0.0998460054397583
step: 190, loss: 0.008814016357064247
step: 200, loss: 0.07248570770025253
step: 210, loss: 0.0313209630548954
/usr/local/lib/python3.7/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=0.6769759450171822, f1=0.6907216494845361,
best_f1=0.7061143984220907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04885415360331535
step: 10, loss: 0.11873279511928558
step: 20, loss: 0.0702778622508049

```

```

step: 30, loss: 0.027824806049466133
step: 40, loss: 0.040686119347810745
step: 50, loss: 0.03078874573111534
step: 60, loss: 0.21042872965335846
step: 70, loss: 0.02509266883134842
step: 80, loss: 0.13838598132133484
step: 90, loss: 0.17371398210525513
step: 100, loss: 0.011687690392136574
step: 110, loss: 0.03130688518285751
step: 120, loss: 0.07034594565629959
step: 130, loss: 0.13652639091014862
step: 140, loss: 0.12984582781791687
step: 150, loss: 0.018820632249116898
step: 160, loss: 0.008530314080417156
step: 170, loss: 0.11794750392436981
step: 180, loss: 0.08339189738035202
step: 190, loss: 0.13015815615653992
step: 200, loss: 0.0068474761210381985
step: 210, loss: 0.15008221566677094
/usr/local/lib/python3.7/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=0.6777546777546777, f1=0.6983471074380165,
best_f1=0.7061143984220907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011027798987925053
step: 10, loss: 0.020215462893247604
step: 20, loss: 0.07649460434913635
step: 30, loss: 0.06926140934228897
step: 40, loss: 0.06461980193853378
step: 50, loss: 0.12333637475967407
step: 60, loss: 0.034631356596946716
step: 70, loss: 0.010826396755874157
step: 80, loss: 0.13327106833457947
step: 90, loss: 0.12830764055252075
step: 100, loss: 0.009526758454740047
step: 110, loss: 0.16844260692596436
step: 120, loss: 0.1097707599401474
step: 130, loss: 0.046863164752721786
step: 140, loss: 0.014703159220516682
step: 150, loss: 0.009566185064613819
step: 160, loss: 0.08242816478013992
step: 170, loss: 0.029141515493392944
step: 180, loss: 0.038540277630090714
step: 190, loss: 0.10413594543933868
step: 200, loss: 0.010507565923035145

```

```

step: 210, loss: 0.042965952306985855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6923076923076924, f1=0.701252236135957,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07410670816898346
step: 10, loss: 0.19277197122573853
step: 20, loss: 0.02156531810760498
step: 30, loss: 0.03132588788866997
step: 40, loss: 0.020703814923763275
step: 50, loss: 0.020637931302189827
step: 60, loss: 0.39070528745651245
step: 70, loss: 0.05711817368865013
step: 80, loss: 0.038650039583444595
step: 90, loss: 0.026027848944067955
step: 100, loss: 0.14324913918972015
step: 110, loss: 0.04355597496032715
step: 120, loss: 0.006907642353326082
step: 130, loss: 0.005282598081976175
step: 140, loss: 0.08912045508623123
step: 150, loss: 0.013082848861813545
step: 160, loss: 0.031193826347589493
step: 170, loss: 0.015054505318403244
step: 180, loss: 0.0586952343583107
step: 190, loss: 0.037973321974277496
step: 200, loss: 0.011167226359248161
step: 210, loss: 0.286761999130249
/usr/local/lib/python3.7/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=0.6778398510242086, f1=0.6982922201138521,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006167857442051172
step: 10, loss: 0.034156836569309235
step: 20, loss: 0.0019791019149124622
step: 30, loss: 0.04115656763315201
step: 40, loss: 0.04100629314780235

```

```

step: 50, loss: 0.20135027170181274
step: 60, loss: 0.07764881104230881
step: 70, loss: 0.07576722651720047
step: 80, loss: 0.029520440846681595
step: 90, loss: 0.003421330824494362
step: 100, loss: 0.0025587992276996374
step: 110, loss: 0.004936751443892717
step: 120, loss: 0.006509595550596714
step: 130, loss: 0.010361958295106888
step: 140, loss: 0.05021631345152855
step: 150, loss: 0.125501349568367
step: 160, loss: 0.0044719805009663105
step: 170, loss: 0.05306931212544441
step: 180, loss: 0.07973000407218933
step: 190, loss: 0.0006572895799763501
step: 200, loss: 0.1395018845796585
step: 210, loss: 0.02327364683151245
/usr/local/lib/python3.7/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=0.6850393700787402, f1=0.6952965235173825,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03439347445964813
step: 10, loss: 0.06347015500068665
step: 20, loss: 0.0005900156684219837
step: 30, loss: 0.0666542574763298
step: 40, loss: 0.00025947910035029054
step: 50, loss: 0.01429730374366045
step: 60, loss: 0.04071634262800217
step: 70, loss: 0.1369975358247757
step: 80, loss: 0.026393912732601166
step: 90, loss: 0.060459259897470474
step: 100, loss: 0.005527493078261614
step: 110, loss: 0.006975518073886633
step: 120, loss: 0.045264892280101776
step: 130, loss: 0.014377760700881481
step: 140, loss: 0.013527769595384598
step: 150, loss: 0.1907815933227539
step: 160, loss: 0.088844895362854
step: 170, loss: 0.01374605018645525
step: 180, loss: 0.05894077196717262
step: 190, loss: 0.08393289893865585
step: 200, loss: 0.01071153860539198
step: 210, loss: 0.12473832815885544
/usr/local/lib/python3.7/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=0.6844106463878327, f1=0.6885880077369438,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024105165153741837
step: 10, loss: 0.023446425795555115
step: 20, loss: 0.01961345411837101
step: 30, loss: 0.0006189939449541271
step: 40, loss: 0.039782147854566574
step: 50, loss: 0.049242500215768814
step: 60, loss: 0.03431084379553795
step: 70, loss: 0.011812862008810043
step: 80, loss: 0.09869842976331711
step: 90, loss: 0.002642543986439705
step: 100, loss: 0.11283881217241287
step: 110, loss: 0.006304075475782156
step: 120, loss: 0.04735332727432251
step: 130, loss: 0.008182032965123653
step: 140, loss: 0.0015425396850332618
step: 150, loss: 0.0002290531265316531
step: 160, loss: 0.05749979987740517
step: 170, loss: 0.02218776009976864
step: 180, loss: 0.013987434096634388
step: 190, loss: 0.022289596498012543
step: 200, loss: 0.001386145013384521
step: 210, loss: 0.005787124391645193
/usr/local/lib/python3.7/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=0.6852589641434264, f1=0.7089108910891089,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11264386028051376
step: 10, loss: 0.0035758614540100098
step: 20, loss: 0.061788760125637054
step: 30, loss: 0.23574812710285187
step: 40, loss: 0.03136065602302551
step: 50, loss: 0.0009748605079948902
step: 60, loss: 0.020177237689495087
step: 70, loss: 0.006185057573020458
step: 80, loss: 0.004466377664357424
step: 90, loss: 0.06276538968086243
step: 100, loss: 0.02096240036189556

```

```

step: 110, loss: 0.006906007882207632
step: 120, loss: 0.03693541884422302
step: 130, loss: 0.0070320013910532
step: 140, loss: 0.0006661557708866894
step: 150, loss: 0.019683262333273888
step: 160, loss: 0.0035950811579823494
step: 170, loss: 0.0013082987861707807
step: 180, loss: 0.021648500114679337
step: 190, loss: 0.02469494193792343
step: 200, loss: 0.0015181826893240213
step: 210, loss: 0.014087763614952564
/usr/local/lib/python3.7/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=0.6799276672694394, f1=0.7050359712230215,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0339500866830349
step: 10, loss: 0.0014437180943787098
step: 20, loss: 0.0877714455127716
step: 30, loss: 0.10732824355363846
step: 40, loss: 0.10368324816226959
step: 50, loss: 0.11375215649604797
step: 60, loss: 0.00729290209710598
step: 70, loss: 0.0521562322974205
step: 80, loss: 0.003363135503605008
step: 90, loss: 0.0012521188473328948
step: 100, loss: 0.01795264519751072
step: 110, loss: 0.003991645760834217
step: 120, loss: 0.044787488877773285
step: 130, loss: 0.1021079421043396
step: 140, loss: 0.024150943383574486
step: 150, loss: 0.02384251356124878
step: 160, loss: 0.011486591771245003
step: 170, loss: 0.007697469089180231
step: 180, loss: 0.050967421382665634
step: 190, loss: 0.001197763136588037
step: 200, loss: 0.10827844589948654
step: 210, loss: 0.01428291667252779
/usr/local/lib/python3.7/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=0.6767485822306237, f1=0.7007575757575758,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004699441138654947
step: 10, loss: 0.03164977952837944
step: 20, loss: 0.023694513365626335
step: 30, loss: 0.007242139894515276
step: 40, loss: 0.00399715406820178
step: 50, loss: 0.0065675461664795876
step: 60, loss: 0.010937320999801159
step: 70, loss: 0.005880057346075773
step: 80, loss: 0.03976675868034363
step: 90, loss: 0.28234657645225525
step: 100, loss: 0.08738303184509277
step: 110, loss: 0.003681976580992341
step: 120, loss: 0.042272135615348816
step: 130, loss: 0.0027789694722741842
step: 140, loss: 0.11288373917341232
step: 150, loss: 0.0016947799595072865
step: 160, loss: 0.004246569238603115
step: 170, loss: 0.006251694168895483
step: 180, loss: 0.0013272492215037346
step: 190, loss: 0.0031918238382786512
step: 200, loss: 0.03925537317991257
step: 210, loss: 0.0055258991196751595
/usr/local/lib/python3.7/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=0.6804511278195489, f1=0.6918714555765595,
best_f1=0.701252236135957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00228852778673172
step: 10, loss: 0.0016689160838723183
step: 20, loss: 0.015004631131887436
step: 30, loss: 0.038993701338768005
step: 40, loss: 0.010362846776843071
step: 50, loss: 0.001087747048586607
step: 60, loss: 0.08046828955411911
step: 70, loss: 0.0017919668462127447
step: 80, loss: 0.006400351878255606
step: 90, loss: 0.010985152795910835
step: 100, loss: 0.00115242472384125
step: 110, loss: 0.09270019084215164
step: 120, loss: 0.0129461120814085
step: 130, loss: 0.004438009113073349
step: 140, loss: 0.0004340244340710342
step: 150, loss: 0.0008912024204619229
step: 160, loss: 0.007952350191771984

```

```

step: 170, loss: 0.0017658602446317673
step: 180, loss: 0.003026730613783002
step: 190, loss: 0.004171648062765598
step: 200, loss: 0.013297222554683685
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 0.01999894343316555
/usr/local/lib/python3.7/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=0.6830188679245283, f1=0.6933333333333335,
best_f1=0.701252236135957
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 520.12it/s]
load_f1 = 0.6808510638297872
real_f1 = 0.6768642447418738
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 262.25it/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: 0.5699781775474548
step: 10, loss: 0.35451167821884155
step: 20, loss: 0.2838574945926666
step: 30, loss: 0.4453263282775879
step: 40, loss: 0.42904043197631836
step: 50, loss: 0.2980559766292572
step: 60, loss: 0.29620155692100525
step: 70, loss: 0.24664075672626495
step: 80, loss: 0.18129189312458038
step: 90, loss: 0.25291892886161804
step: 100, loss: 0.3365437984466553
step: 110, loss: 0.3833475410938263
step: 120, loss: 0.15163178741931915
step: 130, loss: 0.11450318247079849
step: 140, loss: 0.09682172536849976
step: 150, loss: 0.1492835283279419
step: 160, loss: 0.10805322229862213
step: 170, loss: 0.39696118235588074
step: 180, loss: 0.026090558618307114
step: 190, loss: 0.18019431829452515
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7123287671232876, f1=0.7479224376731302,
best_f1=0.7479224376731302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16936391592025757
step: 10, loss: 0.07660941034555435
step: 20, loss: 0.04710846394300461
step: 30, loss: 0.11257514357566833
step: 40, loss: 0.2180902063846588
step: 50, loss: 0.03339379280805588
step: 60, loss: 0.1893939971923828
step: 70, loss: 0.15261195600032806
step: 80, loss: 0.21677935123443604
step: 90, loss: 0.09650437533855438
step: 100, loss: 0.017788151279091835
step: 110, loss: 0.032428111881017685
step: 120, loss: 0.2628611922264099
step: 130, loss: 0.058543648570775986
step: 140, loss: 0.04409833997488022
step: 150, loss: 0.05562556907534599
step: 160, loss: 0.04547356814146042
step: 170, loss: 0.07786048948764801
step: 180, loss: 0.04439372196793556
step: 190, loss: 0.04910372197628021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7197943444730078, f1=0.763819095477387,
best_f1=0.763819095477387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05610227957367897
step: 10, loss: 0.25222840905189514
step: 20, loss: 0.05651816725730896
step: 30, loss: 0.04777327924966812
step: 40, loss: 0.0795883759856224
step: 50, loss: 0.154994934797287
step: 60, loss: 0.02048228308558464
step: 70, loss: 0.18024824559688568
step: 80, loss: 0.04249071702361107
step: 90, loss: 0.023673783987760544

```

```

step: 100, loss: 0.02712329663336277
step: 110, loss: 0.09974958747625351
step: 120, loss: 0.05038634315133095
step: 130, loss: 0.0110911400988698
step: 140, loss: 0.02269025333225727
step: 150, loss: 0.11284570395946503
step: 160, loss: 0.07004359364509583
step: 170, loss: 0.157958522439003
step: 180, loss: 0.03732016682624817
step: 190, loss: 0.13795821368694305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8159203980099502, f1=0.8039215686274511,
best_f1=0.8039215686274511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03168352693319321
step: 10, loss: 0.03464774042367935
step: 20, loss: 0.042428601533174515
step: 30, loss: 0.020271753892302513
step: 40, loss: 0.011494891718029976
step: 50, loss: 0.12035927176475525
step: 60, loss: 0.002226099604740739
step: 70, loss: 0.025314856320619583
step: 80, loss: 0.02788296528160572
step: 90, loss: 0.013413166627287865
step: 100, loss: 0.06112728640437126
step: 110, loss: 0.0048986212350428104
step: 120, loss: 0.10825258493423462
step: 130, loss: 0.16913430392742157
step: 140, loss: 0.0210245493799448
step: 150, loss: 0.15081848204135895
step: 160, loss: 0.005315171089023352
step: 170, loss: 0.08153155446052551
step: 180, loss: 0.03286893665790558
step: 190, loss: 0.1053621843457222
/usr/local/lib/python3.7/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=0.8041237113402062, f1=0.8144329896907216,
best_f1=0.8039215686274511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.061078477650880814
step: 10, loss: 0.011827841401100159
step: 20, loss: 0.02352508157491684
step: 30, loss: 0.004764865152537823
step: 40, loss: 0.05016550421714783
step: 50, loss: 0.1444157063961029
step: 60, loss: 0.19236011803150177
step: 70, loss: 0.005877800285816193
step: 80, loss: 0.02766936831176281
step: 90, loss: 0.014568551443517208
step: 100, loss: 0.005263955798000097
step: 110, loss: 0.03414417803287506
step: 120, loss: 0.07289022952318192
step: 130, loss: 0.03239649161696434
step: 140, loss: 0.2029935121536255
step: 150, loss: 0.10541223734617233
step: 160, loss: 0.006598261184990406
step: 170, loss: 0.0075892130844295025
step: 180, loss: 0.0036822722759097815
step: 190, loss: 0.2200898975133896
```

```
/usr/local/lib/python3.7/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=0.7794871794871795, f1=0.7772020725388601,
best_f1=0.8039215686274511
```

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

An input tensor was not cuda.

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

```
step: 0, loss: 0.004638497717678547
step: 10, loss: 0.16063757240772247
step: 20, loss: 0.0036654204595834017
step: 30, loss: 0.0007126442505978048
step: 40, loss: 0.16200925409793854
step: 50, loss: 0.08330821990966797
step: 60, loss: 0.056572046130895615
step: 70, loss: 0.004827443510293961
step: 80, loss: 0.05534499138593674
step: 90, loss: 0.010083210654556751
step: 100, loss: 0.0011031923349946737
step: 110, loss: 0.0136131402105093
step: 120, loss: 0.007056171074509621
step: 130, loss: 0.0011086029699072242
step: 140, loss: 0.0694580003619194
step: 150, loss: 0.10217422246932983
step: 160, loss: 0.19296258687973022
step: 170, loss: 0.03993425890803337
```

```

step: 180, loss: 0.1314888596534729
step: 190, loss: 0.03257502615451813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8392370572207084, f1=0.827027027027027,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002752625150606036
step: 10, loss: 0.00979689136147499
step: 20, loss: 0.05158174782991409
step: 30, loss: 0.08547916263341904
step: 40, loss: 0.003368944860994816
step: 50, loss: 0.0564272366464138
step: 60, loss: 0.0287966076284647
step: 70, loss: 0.006423100363463163
step: 80, loss: 0.006011002231389284
step: 90, loss: 0.03111417219042778
step: 100, loss: 0.00037111941492184997
step: 110, loss: 0.003999999724328518
step: 120, loss: 0.003191048977896571
step: 130, loss: 0.008051341399550438
step: 140, loss: 0.01116455253213644
step: 150, loss: 0.003359375521540642
step: 160, loss: 0.07416082173585892
step: 170, loss: 0.06546732038259506
step: 180, loss: 0.00544802937656641
step: 190, loss: 0.006053868215531111
/usr/local/lib/python3.7/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=0.7978723404255318, f1=0.8054794520547945,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012869290076196194
step: 10, loss: 0.008639111183583736
step: 20, loss: 0.006985308136790991
step: 30, loss: 0.1233537346124649
step: 40, loss: 0.011758564971387386
step: 50, loss: 0.039034634828567505

```

```

step: 60, loss: 0.002529673045501113
step: 70, loss: 0.010630606673657894
step: 80, loss: 0.0003822915896307677
step: 90, loss: 0.0007339597796089947
step: 100, loss: 0.09667719155550003
step: 110, loss: 0.0010144906118512154
step: 120, loss: 0.0019465689547359943
step: 130, loss: 0.004838172812014818
step: 140, loss: 0.06657235324382782
step: 150, loss: 0.0028541681822389364
step: 160, loss: 0.003088679164648056
step: 170, loss: 0.0028332816436886787
step: 180, loss: 0.010018298402428627
step: 190, loss: 0.0012288980651646852
/usr/local/lib/python3.7/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=0.8207792207792207, f1=0.8150134048257373,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009089717059396207
step: 10, loss: 0.041570357978343964
step: 20, loss: 0.0012929786462336779
step: 30, loss: 0.04107542708516121
step: 40, loss: 0.009844313375651836
step: 50, loss: 0.0008349087438546121
step: 60, loss: 0.007925977930426598
step: 70, loss: 0.0009352326160296798
step: 80, loss: 0.0008088716422207654
step: 90, loss: 0.011450307443737984
step: 100, loss: 0.12136194854974747
step: 110, loss: 0.003640548326075077
step: 120, loss: 0.007416656706482172
step: 130, loss: 0.0008810422732494771
step: 140, loss: 0.0014301505871117115
step: 150, loss: 0.0018367278389632702
step: 160, loss: 0.000782106420956552
step: 170, loss: 0.0070782313123345375
step: 180, loss: 0.0443434938788414
step: 190, loss: 0.0005520347622223198
/usr/local/lib/python3.7/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=0.8238341968911919, f1=0.8194070080862533,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025361916050314903
step: 10, loss: 0.0051333107985556126
step: 20, loss: 0.03842443972826004
step: 30, loss: 0.0014362928923219442
step: 40, loss: 0.003800151636824012
step: 50, loss: 0.008211847394704819
step: 60, loss: 0.00112336000893265
step: 70, loss: 0.005975739564746618
step: 80, loss: 0.002476684283465147
step: 90, loss: 0.017612077295780182
step: 100, loss: 0.0012496465351432562
step: 110, loss: 0.0054352483712136745
step: 120, loss: 0.11718540638685226
step: 130, loss: 0.0016137943603098392
step: 140, loss: 0.0010735302930697799
step: 150, loss: 0.00854654423892498
step: 160, loss: 0.0065541137009859085
step: 170, loss: 0.0029374072328209877
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.00069428764982149
step: 190, loss: 0.0006611381541006267
/usr/local/lib/python3.7/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=0.8216216216216217, f1=0.8033240997229917,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002827031712513417
step: 10, loss: 0.0006274540210142732
step: 20, loss: 0.10157447308301926
step: 30, loss: 0.02049615979194641
step: 40, loss: 0.0006557822343893349
step: 50, loss: 0.023941464722156525
step: 60, loss: 0.000495491607580334
step: 70, loss: 0.007310027722269297
step: 80, loss: 0.09125165641307831
step: 90, loss: 0.0005413079052232206
step: 100, loss: 0.0015949568478390574
step: 110, loss: 0.0009627414401620626
step: 120, loss: 0.0006129876710474491
step: 130, loss: 0.0034632766619324684
step: 140, loss: 0.007406965829432011
step: 150, loss: 0.0030831501353532076
step: 160, loss: 0.0009270498994737864

```



```

step: 170, loss: 0.008833931758999825
step: 180, loss: 0.010065481998026371
step: 190, loss: 0.0007095503387972713
/usr/local/lib/python3.7/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=0.8203753351206434, f1=0.8130081300813009,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024841679260134697
step: 10, loss: 0.0015175826847553253
step: 20, loss: 0.0003659161156974733
step: 30, loss: 0.0014575386885553598
step: 40, loss: 0.006956484168767929
step: 50, loss: 0.00037571703433059156
step: 60, loss: 0.00044697709381580353
step: 70, loss: 0.0012107959482818842
step: 80, loss: 0.0009642087970860302
step: 90, loss: 0.0011399841168895364
step: 100, loss: 0.00045305219828151166
step: 110, loss: 0.0009375881054438651
step: 120, loss: 0.000703484402038157
step: 130, loss: 0.0017014886252582073
step: 140, loss: 0.00034868685179390013
step: 150, loss: 0.0008546543540433049
step: 160, loss: 0.0004103299870621413
step: 170, loss: 0.0011221090098842978
step: 180, loss: 0.00025597374769859016
step: 190, loss: 0.0004071111907251179
/usr/local/lib/python3.7/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=0.8164383561643836, f1=0.8119891008174387,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21203726530075073
step: 10, loss: 0.002484370721504092
step: 20, loss: 0.0006662485538981855
step: 30, loss: 0.00038969673914834857
step: 40, loss: 0.002572538098320365
step: 50, loss: 0.005258898250758648
step: 60, loss: 0.001532439491711557
step: 70, loss: 0.002826014533638954
step: 80, loss: 0.0553092323243618

```

```

step: 90, loss: 0.0010988849680870771
step: 100, loss: 0.00043930773972533643
step: 110, loss: 0.05962485074996948
step: 120, loss: 0.0018054109532386065
step: 130, loss: 0.001363886403851211
step: 140, loss: 0.0038193378131836653
step: 150, loss: 0.000429287989391014
step: 160, loss: 0.0007180978427641094
step: 170, loss: 0.0002215323765994981
step: 180, loss: 0.0016587470890954137
step: 190, loss: 0.0007586218416690826
/usr/local/lib/python3.7/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=0.8152173913043478, f1=0.8174386920980926,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008941580890677869
step: 10, loss: 0.0005788073758594692
step: 20, loss: 0.00019673569477163255
step: 30, loss: 0.0003414527454879135
step: 40, loss: 0.0005489079048857093
step: 50, loss: 0.0017515922663733363
step: 60, loss: 0.01911674067378044
step: 70, loss: 0.0004473357112146914
step: 80, loss: 0.0003879512078128755
step: 90, loss: 0.002397300908342004
step: 100, loss: 0.007843914441764355
step: 110, loss: 0.003703809343278408
step: 120, loss: 0.003233546856790781
step: 130, loss: 0.0011076475493609905
step: 140, loss: 0.002362429164350033
step: 150, loss: 0.001486747758463025
step: 160, loss: 0.001609504921361804
step: 170, loss: 0.004745746497064829
step: 180, loss: 0.005150050390511751
step: 190, loss: 0.0031115347519516945
/usr/local/lib/python3.7/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=0.8174386920980926, f1=0.8021978021978022,
best_f1=0.827027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01461763959378004

```

```

step: 10, loss: 0.0005604436155408621
step: 20, loss: 0.0007438682951033115
step: 30, loss: 0.000931335031054914
step: 40, loss: 0.07668157666921616
step: 50, loss: 0.0013270203489810228
step: 60, loss: 0.01465968694537878
step: 70, loss: 0.003343365853652358
step: 80, loss: 0.0015932298265397549
step: 90, loss: 0.0007517896592617035
step: 100, loss: 0.0005465747672133148
step: 110, loss: 0.0020517485681921244
step: 120, loss: 0.0011391635052859783
step: 130, loss: 0.005890950094908476
step: 140, loss: 0.018769748508930206
step: 150, loss: 0.001411867211572826
step: 160, loss: 0.0010475816670805216
step: 170, loss: 0.0016194243216887116
step: 180, loss: 0.07255139201879501
step: 190, loss: 0.002401127479970455
/usr/local/lib/python3.7/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=0.8216216216216217, f1=0.8087431693989071,
best_f1=0.827027027027027
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:08, 240.27it/s]
load_f1 = 0.7005076142131978
real_f1 = 0.6865671641791045
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, 254.92it/s]

```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

```

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

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

```
step: 0, loss: 0.625533401966095  
step: 10, loss: 0.38102951645851135  
step: 20, loss: 0.30588364601135254  
step: 30, loss: 0.39012569189071655  
step: 40, loss: 0.28143781423568726  
step: 50, loss: 0.2668764293193817  
step: 60, loss: 0.2451813668012619  
step: 70, loss: 0.35044994950294495
```

```

step: 80, loss: 0.3279404938220978
step: 90, loss: 0.2154412865638733
step: 100, loss: 0.2309015989303589
step: 110, loss: 0.2690436840057373
step: 120, loss: 0.11561035364866257
step: 130, loss: 0.024659493938088417
step: 140, loss: 0.0783739760518074
step: 150, loss: 0.22714738547801971
step: 160, loss: 0.09630031883716583
step: 170, loss: 0.18852394819259644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.74934036939314, f1=0.712468193384224,
best_f1=0.712468193384224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1752810776233673
step: 10, loss: 0.17438754439353943
step: 20, loss: 0.09991826117038727
step: 30, loss: 0.15803928673267365
step: 40, loss: 0.1304265409708023
step: 50, loss: 0.08861646801233292
step: 60, loss: 0.13452677428722382
step: 70, loss: 0.09083722531795502
step: 80, loss: 0.0404798723757267
step: 90, loss: 0.08842571079730988
step: 100, loss: 0.1659001260995865
step: 110, loss: 0.061918191611766815
step: 120, loss: 0.1298709362745285
step: 130, loss: 0.03801592066884041
step: 140, loss: 0.3155057430267334
step: 150, loss: 0.09905599057674408
step: 160, loss: 0.18856391310691833
step: 170, loss: 0.06804025918245316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7706422018348624, f1=0.7954545454545455,

```

```

best_f1=0.7954545454545455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033724140375852585
step: 10, loss: 0.027934355661273003
step: 20, loss: 0.12980350852012634
step: 30, loss: 0.1396436095237732
step: 40, loss: 0.05532202124595642
step: 50, loss: 0.06698504090309143
step: 60, loss: 0.055991459637880325
step: 70, loss: 0.04031282290816307
step: 80, loss: 0.024859091266989708
step: 90, loss: 0.04755176231265068
step: 100, loss: 0.009870978072285652
step: 110, loss: 0.1291383057832718
step: 120, loss: 0.10317795723676682
step: 130, loss: 0.026668021455407143
step: 140, loss: 0.08264246582984924
step: 150, loss: 0.012683420442044735
step: 160, loss: 0.043414533138275146
step: 170, loss: 0.0569058395922184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8184143222506394, f1=0.8098765432098766,
best_f1=0.8098765432098766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028890534304082394
step: 10, loss: 0.00935615785419941
step: 20, loss: 0.004761248826980591
step: 30, loss: 0.07058669626712799
step: 40, loss: 0.0011412854073569179
step: 50, loss: 0.13170446455478668
step: 60, loss: 0.013390378095209599
step: 70, loss: 0.020749112591147423
step: 80, loss: 0.05778730288147926
step: 90, loss: 0.09377975016832352
step: 100, loss: 0.05203867703676224
step: 110, loss: 0.052160728722810745
step: 120, loss: 0.008705290034413338
step: 130, loss: 0.019981248304247856

```

```

step: 140, loss: 0.1997663378715515
step: 150, loss: 0.13750244677066803
step: 160, loss: 0.04097794368863106
step: 170, loss: 0.0023348741233348846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8257372654155496, f1=0.8134715025906737,
best_f1=0.8134715025906737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05818264186382294
step: 10, loss: 0.017425989732146263
step: 20, loss: 0.015451709739863873
step: 30, loss: 0.0073022060096263885
step: 40, loss: 0.06369233131408691
step: 50, loss: 0.00399076147004962
step: 60, loss: 0.025258982554078102
step: 70, loss: 0.12399975210428238
step: 80, loss: 0.09566177427768707
step: 90, loss: 0.0979519635438919
step: 100, loss: 0.023133615031838417
step: 110, loss: 0.23594266176223755
step: 120, loss: 0.008507008664309978
step: 130, loss: 0.041092030704021454
step: 140, loss: 0.008497015573084354
step: 150, loss: 0.014273151755332947
step: 160, loss: 0.014635663479566574
step: 170, loss: 0.003958260174840689
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8361858190709045, f1=0.8162291169451074,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005368877202272415
step: 10, loss: 0.004509884398430586

```



```

step: 20, loss: 0.009585175663232803
step: 30, loss: 0.0063373250886797905
step: 40, loss: 0.01586437225341797
step: 50, loss: 0.10324843972921371
step: 60, loss: 0.015901638194918633
step: 70, loss: 0.039655908942222595
step: 80, loss: 0.010031424462795258
step: 90, loss: 0.026281243190169334
step: 100, loss: 0.0031355060636997223
step: 110, loss: 0.001228927168995142
step: 120, loss: 0.019121643155813217
step: 130, loss: 0.01771962083876133
step: 140, loss: 0.006178389769047499
step: 150, loss: 0.018445229157805443
step: 160, loss: 0.0943397656083107
step: 170, loss: 0.004526014439761639
/usr/local/lib/python3.7/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=0.8049999999999999, f1=0.7923627684964201,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005023874109610915
step: 10, loss: 0.006260136142373085
step: 20, loss: 0.0052593788132071495
step: 30, loss: 0.0018279921496286988
step: 40, loss: 0.005015814211219549
step: 50, loss: 0.014952966012060642
step: 60, loss: 0.0004491060972213745
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.0003066512872464955
step: 80, loss: 0.06889691203832626
step: 90, loss: 0.00038039186620153487
step: 100, loss: 0.012768607586622238
step: 110, loss: 0.00399060919880867
step: 120, loss: 0.014449107460677624
step: 130, loss: 0.12605716288089752
step: 140, loss: 0.0012502678437158465
step: 150, loss: 0.007015033159404993
step: 160, loss: 0.0003178865590598434
step: 170, loss: 0.027033846825361252
/usr/local/lib/python3.7/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=0.806366047745358, f1=0.824742268041237,
best_f1=0.8162291169451074

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029653729870915413
step: 10, loss: 0.0001727052585920319
step: 20, loss: 0.0011004838161170483
step: 30, loss: 0.012005624361336231
step: 40, loss: 0.00013771424710284919
step: 50, loss: 0.0015555096324533224
step: 60, loss: 0.0025236785877496004
step: 70, loss: 0.0008398584905080497
step: 80, loss: 0.008128136396408081
step: 90, loss: 0.005789367947727442
step: 100, loss: 0.0617499053478241
step: 110, loss: 0.06087207794189453
step: 120, loss: 0.0011116070672869682
step: 130, loss: 0.01631372608244419
step: 140, loss: 0.0005854153423570096
step: 150, loss: 0.04737725853919983
step: 160, loss: 0.024879608303308487
step: 170, loss: 0.0016029342077672482
/usr/local/lib/python3.7/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=0.8267326732673267, f1=0.7990430622009569,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007766312337480485
step: 10, loss: 0.005670546554028988
step: 20, loss: 0.00113869016058743
step: 30, loss: 0.0008373991586267948
step: 40, loss: 0.006858848035335541
step: 50, loss: 0.0005318706971593201
step: 60, loss: 0.004867014475166798
step: 70, loss: 0.01591603457927704
step: 80, loss: 0.00047865620581433177
step: 90, loss: 0.03756054490804672
step: 100, loss: 0.016680896282196045
step: 110, loss: 0.0029037001077085733
step: 120, loss: 0.02944963611662388
step: 130, loss: 0.05637883394956589
step: 140, loss: 0.0036958425771445036
step: 150, loss: 0.0004361277096904814
step: 160, loss: 0.0035471494775265455
step: 170, loss: 0.00045348150888457894
/usr/local/lib/python3.7/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=0.8337730870712401, f1=0.837092731829574,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032450444996356964
step: 10, loss: 0.0010671511990949512
step: 20, loss: 0.07135891914367676
step: 30, loss: 0.0006961998878978193
step: 40, loss: 0.00010211637709289789
step: 50, loss: 0.0555332750082016
step: 60, loss: 0.00012055144179612398
step: 70, loss: 0.001338283414952457
step: 80, loss: 0.0001605491415830329
step: 90, loss: 0.011180147528648376
step: 100, loss: 0.00019399909069761634
step: 110, loss: 0.0022330889478325844
step: 120, loss: 0.0012229442363604903
step: 130, loss: 0.0012558003654703498
step: 140, loss: 0.05718858167529106
step: 150, loss: 0.01361576747149229
step: 160, loss: 0.006216020323336124
step: 170, loss: 0.001984038855880499
/usr/local/lib/python3.7/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=0.8210526315789473, f1=0.8241206030150753,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04279926046729088
step: 10, loss: 0.04214818403124809
step: 20, loss: 0.0015095340786501765
step: 30, loss: 0.0001253182563232258
step: 40, loss: 0.00048544714809395373
step: 50, loss: 0.014699988067150116
step: 60, loss: 0.004942683503031731
step: 70, loss: 0.00011807090049842373
step: 80, loss: 0.0001716033584671095
step: 90, loss: 0.0007960480870679021
step: 100, loss: 0.00016789877554401755
step: 110, loss: 0.11484108120203018
step: 120, loss: 0.00029931103927083313
step: 130, loss: 9.859416604740545e-05
step: 140, loss: 0.02641715295612812

```

```

step: 150, loss: 0.0011882407125085592
step: 160, loss: 0.0032566434238106012
step: 170, loss: 0.0005754082230851054
/usr/local/lib/python3.7/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=0.8170426065162907, f1=0.8056206088992974,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029816916212439537
step: 10, loss: 0.00035596167435869575
step: 20, loss: 0.0006147893727757037
step: 30, loss: 0.00017322292842436582
step: 40, loss: 0.00010160023521166295
step: 50, loss: 6.327516166493297e-05
step: 60, loss: 0.0018103517359122634
step: 70, loss: 0.057649385184049606
step: 80, loss: 8.144124876707792e-05
step: 90, loss: 0.001256907475180924
step: 100, loss: 0.014476786367595196
step: 110, loss: 0.0001606914447620511
step: 120, loss: 0.0014865221455693245
step: 130, loss: 0.005604192148894072
step: 140, loss: 0.001020443974994123
step: 150, loss: 0.002553186845034361
step: 160, loss: 0.00011619243741733953
step: 170, loss: 0.00012744456762447953
/usr/local/lib/python3.7/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=0.7741935483870969, f1=0.8020304568527918,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016788724809885025
step: 10, loss: 0.00012987118680030107
step: 20, loss: 9.025048348121345e-05
step: 30, loss: 7.21874093869701e-05
step: 40, loss: 0.00033675660961307585
step: 50, loss: 9.623968071537092e-05
step: 60, loss: 0.002418232848867774
step: 70, loss: 0.0016276941169053316
step: 80, loss: 0.0004702743608504534
step: 90, loss: 7.078883936628699e-05
step: 100, loss: 0.0016885182121768594

```

```

step: 110, loss: 6.03503649472259e-05
step: 120, loss: 0.038470372557640076
step: 130, loss: 6.386284076143056e-05
step: 140, loss: 6.886439950903878e-05
step: 150, loss: 0.009933036752045155
step: 160, loss: 0.0009257494239136577
step: 170, loss: 0.0008845457341521978
/usr/local/lib/python3.7/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=0.8102564102564102, f1=0.8329297820823245,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015973453992046416
step: 10, loss: 5.795466131530702e-05
step: 20, loss: 7.438027387252077e-05
step: 30, loss: 0.00045760226203128695
step: 40, loss: 8.014530612854287e-05
step: 50, loss: 0.00011058766540372744
step: 60, loss: 0.0018529290100559592
step: 70, loss: 0.00012818846153095365
step: 80, loss: 0.0012417795369401574
step: 90, loss: 0.0005216123536229134
step: 100, loss: 0.0005350179853849113
step: 110, loss: 8.384909597225487e-05
step: 120, loss: 0.08168601989746094
step: 130, loss: 0.010746914893388748
step: 140, loss: 6.694690091535449e-05
step: 150, loss: 5.710600089514628e-05
step: 160, loss: 4.975661795469932e-05
step: 170, loss: 0.00938281137496233
/usr/local/lib/python3.7/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=0.8163265306122449, f1=0.8317307692307692,
best_f1=0.8162291169451074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001338558504357934
step: 10, loss: 0.005094087682664394
step: 20, loss: 0.011088012717664242
step: 30, loss: 0.0012134021380916238
step: 40, loss: 0.00018210663984064013
step: 50, loss: 8.046186121646315e-05
step: 60, loss: 0.006755025126039982

```

```

step: 70, loss: 4.9570578994462267e-05
step: 80, loss: 0.014037729240953922
step: 90, loss: 0.00041407515527680516
step: 100, loss: 0.0008027347503229976
step: 110, loss: 7.75604639784433e-05
step: 120, loss: 0.00011964800069108605
step: 130, loss: 0.003077787347137928
step: 140, loss: 7.344628102146089e-05
step: 150, loss: 0.00022171306773088872
step: 160, loss: 9.483840403845534e-05
step: 170, loss: 0.00012912781676277518
/usr/local/lib/python3.7/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=0.8174386920980926, f1=0.8320413436692506,
best_f1=0.8162291169451074
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, 266.09it/s]
load_f1 = 0.29950900163666117

```

```

real_f1 = 0.28837209302325584
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:17, 258.12it/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, 434kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.49MB/s]
Downloading: 100% 440M/440M [00:13<00:00, 33.7MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6375412940979004
step: 10, loss: 0.6177398562431335
step: 20, loss: 0.34763652086257935
step: 30, loss: 0.14155258238315582
step: 40, loss: 0.23479777574539185
step: 50, loss: 0.051354262977838516
step: 60, loss: 0.09106528013944626
step: 70, loss: 0.06333059072494507
step: 80, loss: 0.04697385057806969
step: 90, loss: 0.12588830292224884
step: 100, loss: 0.00859972182661295
step: 110, loss: 0.26929306983947754
step: 120, loss: 0.007101788651198149
step: 130, loss: 0.01706446334719658

```



```

step: 140, loss: 0.00206519803032279
step: 150, loss: 0.01650339737534523
step: 160, loss: 0.004348736256361008
step: 170, loss: 0.11022403836250305
step: 180, loss: 0.021264512091875076
step: 190, loss: 0.0727352574467659
step: 200, loss: 0.059106189757585526
step: 210, loss: 0.015548701398074627
step: 220, loss: 0.014263251796364784
step: 230, loss: 0.011963610537350178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9766925638179801, 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: 0.00504769803956151
step: 10, loss: 0.0015481467125937343
step: 20, loss: 0.15596914291381836
step: 30, loss: 0.19365237653255463
step: 40, loss: 0.13783800601959229
step: 50, loss: 0.0036292336881160736
step: 60, loss: 0.0033430736511945724
step: 70, loss: 0.01609664410352707
step: 80, loss: 0.010937524028122425
step: 90, loss: 0.04914534464478493
step: 100, loss: 0.036912307143211365
step: 110, loss: 0.08233150839805603
step: 120, loss: 0.11349540948867798
step: 130, loss: 0.11080842465162277
step: 140, loss: 0.0018943891627714038
step: 150, loss: 0.04906890168786049
step: 160, loss: 0.060147009789943695
step: 170, loss: 0.0017495949286967516
step: 180, loss: 0.004631916061043739
step: 190, loss: 0.006316429004073143
step: 200, loss: 0.0020673840772360563
step: 210, loss: 0.0021162054035812616
step: 220, loss: 0.07946828752756119
step: 230, loss: 0.013086318969726562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775784753363228, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002679202239960432
step: 10, loss: 0.002467684680595994
step: 20, loss: 0.05061439052224159
step: 30, loss: 0.020602283999323845
step: 40, loss: 0.08094360679388046
step: 50, loss: 0.0026842697989195585
step: 60, loss: 0.002317622536793351
step: 70, loss: 0.0022949576377868652
step: 80, loss: 0.001905467826873064
step: 90, loss: 0.015114882029592991
step: 100, loss: 0.0010539545910432935
step: 110, loss: 0.0013371302047744393
step: 120, loss: 0.019340530037879944
step: 130, loss: 0.00038844611844979227
step: 140, loss: 0.004152047913521528
step: 150, loss: 0.015148953534662724
step: 160, loss: 0.06786900013685226
step: 170, loss: 0.0056714462116360664
step: 180, loss: 0.02817993238568306
step: 190, loss: 0.0020809383131563663
step: 200, loss: 0.001490412512794137
step: 210, loss: 0.02581583522260189
step: 220, loss: 0.00047625345177948475
step: 230, loss: 0.0006406758329831064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843400447427293, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008583078742958605
step: 10, loss: 0.002137022791430354

```

```

step: 20, loss: 0.0003188631380908191
step: 30, loss: 0.00046426980406977236
step: 40, loss: 0.005242538638412952
step: 50, loss: 0.0004740818403661251
step: 60, loss: 0.005657934583723545
step: 70, loss: 0.0012552172411233187
step: 80, loss: 0.009526102803647518
step: 90, loss: 0.0008751412387937307
step: 100, loss: 0.007542809471487999
step: 110, loss: 0.0003739613457582891
step: 120, loss: 0.011312552727758884
step: 130, loss: 0.002633915515616536
step: 140, loss: 0.0002867969742510468
step: 150, loss: 0.048565976321697235
step: 160, loss: 0.004326265770941973
step: 170, loss: 0.04907748103141785
step: 180, loss: 0.0006575471488758922
step: 190, loss: 0.0032974977511912584
step: 200, loss: 0.0009180819033645093
step: 210, loss: 0.15868911147117615
step: 220, loss: 0.0006882027955725789
step: 230, loss: 0.016665257513523102
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9775280898876404,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012760601239278913
step: 10, loss: 0.003683104645460844
step: 20, loss: 0.004785980097949505
step: 30, loss: 0.00018887499754782766
step: 40, loss: 0.0011772002326324582
step: 50, loss: 0.0001790640235412866
step: 60, loss: 0.16368801891803741
step: 70, loss: 0.006120082922279835
step: 80, loss: 0.0004224235890433192
step: 90, loss: 0.0031908887904137373
step: 100, loss: 0.00043060636380687356
step: 110, loss: 0.0005312214489094913
step: 120, loss: 0.00013812925317324698
step: 130, loss: 0.0010877953609451652
step: 140, loss: 0.001842829748056829
step: 150, loss: 0.0019090280402451754
step: 160, loss: 0.00012583052739501
step: 170, loss: 0.007103684823960066

```

```

step: 180, loss: 0.00041150161996483803
step: 190, loss: 0.1252574622631073
step: 200, loss: 0.0010829524835571647
step: 210, loss: 0.005921376869082451
step: 220, loss: 0.0004915756289847195
step: 230, loss: 0.0002483152784407139
/usr/local/lib/python3.7/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=0.9830124575311437, f1=0.9751693002257337,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007088602986186743
step: 10, loss: 0.0005933052161708474
step: 20, loss: 0.0018020231509581208
step: 30, loss: 0.00011913547496078536
step: 40, loss: 0.00016372985555790365
step: 50, loss: 0.00019743124721571803
step: 60, loss: 0.00015851085481699556
step: 70, loss: 0.000409232045058161
step: 80, loss: 0.008099744096398354
step: 90, loss: 0.0001689124619588256
step: 100, loss: 0.1450044959783554
step: 110, loss: 0.0007482410874217749
step: 120, loss: 0.000755961169488728
step: 130, loss: 0.0055688112042844296
step: 140, loss: 0.0015793556813150644
step: 150, loss: 0.0008293220889754593
step: 160, loss: 0.0007449627155438066
step: 170, loss: 0.0002701034536585212
step: 180, loss: 0.008901506662368774
step: 190, loss: 0.0005703989299945533
step: 200, loss: 0.0003387074975762516
step: 210, loss: 0.0015986781800165772
step: 220, loss: 0.00017555459635332227
step: 230, loss: 0.09677961468696594
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9797297297297298,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005041492986492813
step: 10, loss: 0.00013424853386823088

```

```

step: 20, loss: 0.016576973721385002
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.00519093731418252
step: 40, loss: 0.001069281017407775
step: 50, loss: 9.212542499881238e-05
step: 60, loss: 0.004338751547038555
step: 70, loss: 0.022915072739124298
step: 80, loss: 0.00655002286657691
step: 90, loss: 0.000743665499612689
step: 100, loss: 0.000381858873879537
step: 110, loss: 0.00023231495288200676
step: 120, loss: 0.0001537835196359083
step: 130, loss: 0.0003534287679940462
step: 140, loss: 0.00017411529552191496
step: 150, loss: 0.01533601526170969
step: 160, loss: 0.03370538726449013
step: 170, loss: 0.0019978575874119997
step: 180, loss: 0.0025799123104661703
step: 190, loss: 0.00016520638018846512
step: 200, loss: 0.052379146218299866
step: 210, loss: 6.945329369045794e-05
step: 220, loss: 0.0034429070074111223
step: 230, loss: 0.00029392982833087444
/usr/local/lib/python3.7/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=0.9808773903262092, f1=0.9775280898876404,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.218503014883026e-05
step: 10, loss: 0.00264939246699214
step: 20, loss: 0.0003025329497177154
step: 30, loss: 6.257231871131808e-05
step: 40, loss: 0.0009534242562949657
step: 50, loss: 0.00022733642254024744
step: 60, loss: 7.118615030776709e-05
step: 70, loss: 0.00033298812923021615
step: 80, loss: 0.00034948159009218216
step: 90, loss: 0.000140121512231417
step: 100, loss: 0.0006271280581131577
step: 110, loss: 0.026044733822345734
step: 120, loss: 0.1109447181224823
step: 130, loss: 0.00016557658091187477
step: 140, loss: 0.0001778778969310224
step: 150, loss: 7.400176400551572e-05
step: 160, loss: 0.0001441482309019193

```

```

step: 170, loss: 9.628212865209207e-05
step: 180, loss: 0.22840772569179535
step: 190, loss: 0.0036289519630372524
step: 200, loss: 0.0006823100848123431
step: 210, loss: 0.009393269196152687
step: 220, loss: 0.00015806358715053648
step: 230, loss: 0.00010710419883253053
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9731543624161074,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014492172340396792
step: 10, loss: 0.0003255395859014243
step: 20, loss: 0.0002650561509653926
step: 30, loss: 0.00014425563858821988
step: 40, loss: 0.01846867986023426
step: 50, loss: 7.143755647120997e-05
step: 60, loss: 0.0007663371507078409
step: 70, loss: 0.00014235571143217385
step: 80, loss: 9.007535118144006e-05
step: 90, loss: 0.00028363976161926985
step: 100, loss: 0.0001294448011321947
step: 110, loss: 4.9251764721702784e-05
step: 120, loss: 5.320227865013294e-05
step: 130, loss: 6.313180347206071e-05
step: 140, loss: 0.0001227321772603318
step: 150, loss: 0.0007040607160888612
step: 160, loss: 0.0001877482864074409
step: 170, loss: 0.00013320418656803668
step: 180, loss: 0.017479585483670235
step: 190, loss: 7.566231943201274e-05
step: 200, loss: 0.00013369494990911335
step: 210, loss: 0.00012225983664393425
step: 220, loss: 5.8686433476395905e-05
step: 230, loss: 8.43703091959469e-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=0.9831271091113611, f1=0.9763779527559054,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.332816989626735e-05

```

```

step: 10, loss: 4.147533036302775e-05
step: 20, loss: 0.0009617786854505539
step: 30, loss: 0.000116993862320669
step: 40, loss: 5.905486978008412e-05
step: 50, loss: 0.0001617258385522291
step: 60, loss: 0.001020687399432063
step: 70, loss: 0.00021111732348799706
step: 80, loss: 0.00010371030657552183
step: 90, loss: 5.972013241262175e-05
step: 100, loss: 0.0011015153722837567
step: 110, loss: 0.0001500842481618747
step: 120, loss: 0.010032621212303638
step: 130, loss: 0.000114469337859191
step: 140, loss: 0.0233895406126976
step: 150, loss: 0.021527810022234917
step: 160, loss: 7.575454947073013e-05
step: 170, loss: 3.670038495329209e-05
step: 180, loss: 7.939760689623654e-05
step: 190, loss: 0.2180638611316681
step: 200, loss: 0.0001554245682200417
step: 210, loss: 0.0008150098146870732
step: 220, loss: 8.685720240464434e-05
step: 230, loss: 0.0024761029053479433
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9752808988764046,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014749592810403556
step: 10, loss: 0.0006130625260993838
step: 20, loss: 0.002820071531459689
step: 30, loss: 0.0006955974386073649
step: 40, loss: 0.0009162799688056111
step: 50, loss: 0.0023157040122896433
step: 60, loss: 0.0016526413382962346
step: 70, loss: 0.004118556622415781
step: 80, loss: 3.9106493204599246e-05
step: 90, loss: 9.124955249717459e-05
step: 100, loss: 4.88204495923128e-05
step: 110, loss: 4.212277053738944e-05
step: 120, loss: 5.026758663007058e-05
step: 130, loss: 3.7775087548652664e-05
step: 140, loss: 4.208529571769759e-05
step: 150, loss: 0.031031515449285507
step: 160, loss: 9.386092278873548e-05

```

```

step: 170, loss: 0.018009228631854057
step: 180, loss: 0.0005077747046016157
step: 190, loss: 0.0002004119596676901
step: 200, loss: 0.0023211361840367317
step: 210, loss: 6.182804645504802e-05
step: 220, loss: 0.00013497736654244363
step: 230, loss: 0.0001375962601741776
/usr/local/lib/python3.7/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=0.980963045912654, f1=0.9753914988814317,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.454292790498585e-05
step: 10, loss: 3.7657915527233854e-05
step: 20, loss: 4.314438410801813e-05
step: 30, loss: 8.483613055432215e-05
step: 40, loss: 0.00015901256119832397
step: 50, loss: 0.005305655766278505
step: 60, loss: 0.00025184612604789436
step: 70, loss: 7.539171201642603e-05
step: 80, loss: 0.0004906005342490971
step: 90, loss: 4.486161196837202e-05
step: 100, loss: 4.958435602020472e-05
step: 110, loss: 4.657152385334484e-05
step: 120, loss: 5.4306929087033495e-05
step: 130, loss: 6.438770651584491e-05
step: 140, loss: 8.069816249189898e-05
step: 150, loss: 0.0009469650103710592
step: 160, loss: 4.5944081648485735e-05
step: 170, loss: 6.889207725180313e-05
step: 180, loss: 0.12142714858055115
step: 190, loss: 0.00036416505463421345
step: 200, loss: 2.8184789698570967e-05
step: 210, loss: 6.425230822060257e-05
step: 220, loss: 3.5258497518952936e-05
step: 230, loss: 0.041013456881046295
/usr/local/lib/python3.7/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=0.9831649831649831, f1=0.9742441209406495,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016460279584862292

```



```

step: 10, loss: 0.0019253389909863472
step: 20, loss: 0.0024742353707551956
step: 30, loss: 0.00019946186512243003
step: 40, loss: 0.00011727953824447468
step: 50, loss: 0.0002186967758461833
step: 60, loss: 0.00034749999758787453
step: 70, loss: 6.334523641271517e-05
step: 80, loss: 5.6407978263450786e-05
step: 90, loss: 0.00018590448598843068
step: 100, loss: 0.0001466104731662199
step: 110, loss: 0.02713003382086754
step: 120, loss: 0.031205182895064354
step: 130, loss: 5.139845598023385e-05
step: 140, loss: 4.6705416025361046e-05
step: 150, loss: 5.802727537229657e-05
step: 160, loss: 0.00020817130280192941
step: 170, loss: 0.00012929114745929837
step: 180, loss: 4.0302144043380395e-05
step: 190, loss: 7.334497786359861e-05
step: 200, loss: 8.091331255855039e-05
step: 210, loss: 3.142175773973577e-05
step: 220, loss: 0.00017681739700492471
step: 230, loss: 0.00020207067427691072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9820224719101124, f1=0.9763779527559054,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.366588786477223e-05
step: 10, loss: 9.99093972495757e-05
step: 20, loss: 8.676491415826604e-05
step: 30, loss: 5.455296559375711e-05
step: 40, loss: 0.014331516809761524
step: 50, loss: 0.00011645485210465267
step: 60, loss: 3.9299509808188304e-05
step: 70, loss: 4.993402399122715e-05
step: 80, loss: 0.00011282296327408403
step: 90, loss: 6.762109114788473e-05
step: 100, loss: 5.444304770207964e-05
step: 110, loss: 9.90616244962439e-05
step: 120, loss: 3.5693898098543286e-05
step: 130, loss: 3.6991204979131e-05
step: 140, loss: 4.7803008783375844e-05
step: 150, loss: 7.115191692719236e-05
step: 160, loss: 0.0033185353968292475

```

```

step: 170, loss: 0.00032015182659961283
step: 180, loss: 0.00019088128465227783
step: 190, loss: 0.00018266485130880028
step: 200, loss: 6.851329089840874e-05
step: 210, loss: 0.0004520573129411787
step: 220, loss: 0.00016436386795248836
step: 230, loss: 0.0015393500216305256
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9753363228699552,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.395870953681879e-05
step: 10, loss: 2.917489291576203e-05
step: 20, loss: 4.488717968342826e-05
step: 30, loss: 0.00011545133020263165
step: 40, loss: 0.00016480643535032868
step: 50, loss: 0.028043657541275024
step: 60, loss: 6.62930469843559e-05
step: 70, loss: 0.00015004098531790078
step: 80, loss: 0.000100878176453989
step: 90, loss: 0.0006129587418399751
step: 100, loss: 3.920695235137828e-05
step: 110, loss: 4.1507460991851985e-05
step: 120, loss: 4.7438006731681526e-05
step: 130, loss: 0.0005487421876750886
step: 140, loss: 3.4931137633975595e-05
step: 150, loss: 0.00010394897981313989
step: 160, loss: 3.5663841117639095e-05
step: 170, loss: 2.719753138080705e-05
step: 180, loss: 0.00011915234063053504
step: 190, loss: 0.00019829111988656223
step: 200, loss: 7.393882697215304e-05
step: 210, loss: 0.0003674036997836083
step: 220, loss: 0.0003390558995306492
step: 230, loss: 3.603357254178263e-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=0.9810055865921787, f1=0.9732142857142857,
best_f1=0.9764837625979844
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, 192.34it/s]
load_f1 = 0.9832402234636871
real_f1 = 0.9821428571428571
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, 239.79it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

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

step: 0, loss: 0.6213669180870056
step: 10, loss: 0.5239959359169006
step: 20, loss: 0.45742180943489075
step: 30, loss: 0.0799039974808693
step: 40, loss: 0.18145491182804108
step: 50, loss: 0.13337965309619904
step: 60, loss: 0.03789757937192917
step: 70, loss: 0.0703146755695343
step: 80, loss: 0.03450696915388107
step: 90, loss: 0.06054216995835304
step: 100, loss: 0.006133676040917635
step: 110, loss: 0.02055242285132408
step: 120, loss: 0.11292842030525208
step: 130, loss: 0.12579117715358734
step: 140, loss: 0.12091178447008133
step: 150, loss: 0.02900487743318081
step: 160, loss: 0.028705701231956482
step: 170, loss: 0.17649586498737335
step: 180, loss: 0.05658762902021408
step: 190, loss: 0.018152011558413506
step: 200, loss: 0.15096226334571838
step: 210, loss: 0.0582759827375412
step: 220, loss: 0.18018688261508942
step: 230, loss: 0.1609678566455841
step: 240, loss: 0.059996601194143295
step: 250, loss: 0.03928312286734581
step: 260, loss: 0.0420445092022419
step: 270, loss: 0.05145987123250961
step: 280, loss: 0.06588220596313477
step: 290, loss: 0.08095261454582214
step: 300, loss: 0.06741220504045486
step: 310, loss: 0.14489030838012695
step: 320, loss: 0.12987834215164185
step: 330, loss: 0.02487405575811863
step: 340, loss: 0.03965624421834946
step: 350, loss: 0.05852924659848213
step: 360, loss: 0.024325495585799217
step: 370, loss: 0.1050945296883583

```

step: 380, loss: 0.023918839171528816
step: 390, loss: 0.05526380613446236
step: 400, loss: 0.37594276666641235
step: 410, loss: 0.03352304920554161
step: 420, loss: 0.18650174140930176
step: 430, loss: 0.15549179911613464
step: 440, loss: 0.03675629198551178
step: 450, loss: 0.003338622860610485
step: 460, loss: 0.03794901818037033
step: 470, loss: 0.061687249690294266
step: 480, loss: 0.04312071204185486
step: 490, loss: 0.0720655620098114
step: 500, loss: 0.0480138324201107
step: 510, loss: 0.056699007749557495
step: 520, loss: 0.07120106369256973
step: 530, loss: 0.008849972859025002
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9387755102040817, f1=0.9342592592592593,
best_f1=0.9342592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04393058270215988
step: 10, loss: 0.009104270488023758
step: 20, loss: 0.021091634407639503
step: 30, loss: 0.007145833224058151
step: 40, loss: 0.06136338412761688
step: 50, loss: 0.16115714609622955
step: 60, loss: 0.03706211596727371
step: 70, loss: 0.02518046461045742
step: 80, loss: 0.06216045096516609
step: 90, loss: 0.012100888416171074
step: 100, loss: 0.02988375350832939
step: 110, loss: 0.026769667863845825
step: 120, loss: 0.036667708307504654
step: 130, loss: 0.17248418927192688
step: 140, loss: 0.02902943082153797
step: 150, loss: 0.1355033963918686
step: 160, loss: 0.004519328940659761
step: 170, loss: 0.02352152392268181
step: 180, loss: 0.02576281875371933
step: 190, loss: 0.05398111417889595

```

```

step: 200, loss: 0.010036596097052097
step: 210, loss: 0.07124462723731995
step: 220, loss: 0.11363769322633743
step: 230, loss: 0.016897421330213547
step: 240, loss: 0.13528631627559662
step: 250, loss: 0.05839941278100014
step: 260, loss: 0.003262563142925501
step: 270, loss: 0.08982880413532257
step: 280, loss: 0.011631677858531475
step: 290, loss: 0.012693325988948345
step: 300, loss: 0.18165741860866547
step: 310, loss: 0.008674060925841331
step: 320, loss: 0.15994204580783844
step: 330, loss: 0.043000344187021255
step: 340, loss: 0.008372999727725983
step: 350, loss: 0.0005034155328758061
step: 360, loss: 0.09643910080194473
step: 370, loss: 0.07410974055528641
step: 380, loss: 0.04414491727948189
step: 390, loss: 0.030662143602967262
step: 400, loss: 0.03195321932435036
step: 410, loss: 0.10253135859966278
step: 420, loss: 0.10095500946044922
step: 430, loss: 0.018774718046188354
step: 440, loss: 0.18978267908096313
step: 450, loss: 0.014589333906769753
step: 460, loss: 0.07110276073217392
step: 470, loss: 0.0640767365694046
step: 480, loss: 0.3022926151752472
step: 490, loss: 0.007984431460499763
step: 500, loss: 0.18497660756111145
step: 510, loss: 0.025492621585726738
step: 520, loss: 0.09084165096282959
step: 530, loss: 0.01991255208849907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9471221338324755, f1=0.9439555349698935,
best_f1=0.9439555349698935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.056076619774103165
step: 10, loss: 0.033570945262908936

```

step: 20, loss: 0.15075789391994476
step: 30, loss: 0.09991797804832458
step: 40, loss: 0.017191607505083084
step: 50, loss: 0.11522377282381058
step: 60, loss: 0.006715542171150446
step: 70, loss: 0.01634451188147068
step: 80, loss: 0.0017973430221900344
step: 90, loss: 0.008102072402834892
step: 100, loss: 0.028107721358537674
step: 110, loss: 0.031470123678445816
step: 120, loss: 0.06454667448997498
step: 130, loss: 0.0034058133605867624
step: 140, loss: 0.11778076738119125
step: 150, loss: 0.15303851664066315
step: 160, loss: 0.02080969326198101
step: 170, loss: 0.10880950093269348
step: 180, loss: 0.015236667357385159
step: 190, loss: 0.00414177356287837
step: 200, loss: 0.0627591609954834
step: 210, loss: 0.11332331597805023
step: 220, loss: 0.07505596429109573
step: 230, loss: 0.09902559220790863
step: 240, loss: 0.0032544410787522793
step: 250, loss: 0.033808380365371704
step: 260, loss: 0.017237789928913116
step: 270, loss: 0.0023068946320563555
step: 280, loss: 0.15591537952423096
step: 290, loss: 0.0006991925765760243
step: 300, loss: 0.019710315391421318
step: 310, loss: 0.009495005942881107
step: 320, loss: 0.013148624449968338
step: 330, loss: 0.009227761067450047
step: 340, loss: 0.002906292909756303
step: 350, loss: 0.001974188955500722
step: 360, loss: 0.039726078510284424
step: 370, loss: 0.03211876377463341
step: 380, loss: 0.01736615039408207
step: 390, loss: 0.005727075971662998
step: 400, loss: 0.0277258288115263
step: 410, loss: 0.006260129623115063
step: 420, loss: 0.08393295854330063
step: 430, loss: 0.006764698773622513
step: 440, loss: 0.0014206254854798317
step: 450, loss: 0.0685054361820221
step: 460, loss: 0.06368743628263474
step: 470, loss: 0.10803284496068954
step: 480, loss: 0.005841712933033705
step: 490, loss: 0.008418395183980465


```

step: 500, loss: 0.003300861455500126
step: 510, loss: 0.02192012593150139
step: 520, loss: 0.088412806391716
step: 530, loss: 0.12209957093000412
/usr/local/lib/python3.7/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=0.9465930018416207, f1=0.938134810710988,
best_f1=0.9439555349698935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0541391521692276
step: 10, loss: 0.008070264011621475
step: 20, loss: 0.0013013302814215422
step: 30, loss: 0.0027146877255290747
step: 40, loss: 0.0009500494343228638
step: 50, loss: 0.12986302375793457
step: 60, loss: 0.061677634716033936
step: 70, loss: 0.0032472943421453238
step: 80, loss: 0.01164490170776844
step: 90, loss: 0.03428139537572861
step: 100, loss: 0.0023263581097126007
step: 110, loss: 0.02478322573006153
step: 120, loss: 0.00020695687271654606
step: 130, loss: 0.004769506864249706
step: 140, loss: 0.018904943019151688
step: 150, loss: 0.02567681297659874
step: 160, loss: 0.10412318259477615
step: 170, loss: 0.004612931981682777
step: 180, loss: 0.005844332743436098
step: 190, loss: 0.004869855474680662
step: 200, loss: 0.0019174475455656648
step: 210, loss: 0.02782856486737728
step: 220, loss: 0.009227007627487183
step: 230, loss: 0.06486526876688004
step: 240, loss: 0.002136067720130086
step: 250, loss: 0.012995989993214607
step: 260, loss: 0.005667630583047867
step: 270, loss: 0.033072397112846375
step: 280, loss: 0.07873179763555527
step: 290, loss: 0.019242607057094574
step: 300, loss: 0.000530554389115423
step: 310, loss: 0.014798819087445736
step: 320, loss: 0.01340070553123951
step: 330, loss: 0.05407293140888214
step: 340, loss: 0.03357606381177902
step: 350, loss: 0.020105551928281784

```

```

step: 360, loss: 0.026309777051210403
step: 370, loss: 0.003505975939333439
step: 380, loss: 0.07665715366601944
step: 390, loss: 0.0034206330310553312
step: 400, loss: 0.024492235854268074
step: 410, loss: 0.006422446109354496
step: 420, loss: 0.0066130440682172775
step: 430, loss: 0.09802430123090744
step: 440, loss: 0.03716178610920906
step: 450, loss: 0.002294834703207016
step: 460, loss: 0.004362780600786209
step: 470, loss: 0.01333875022828579
step: 480, loss: 0.10777430981397629
step: 490, loss: 0.002086981665343046
step: 500, loss: 0.016097595915198326
step: 510, loss: 0.026205625385046005
step: 520, loss: 0.0647052451968193
step: 530, loss: 0.014481723308563232
/usr/local/lib/python3.7/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=0.943778801843318, f1=0.9330889092575618,
best_f1=0.9439555349698935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016376918181777
step: 10, loss: 0.09332647919654846
step: 20, loss: 0.0023735982831567526
step: 30, loss: 0.006218547932803631
step: 40, loss: 0.13004574179649353
step: 50, loss: 0.003896984038874507
step: 60, loss: 0.027289897203445435
step: 70, loss: 0.008040407672524452
step: 80, loss: 0.001925822114571929
step: 90, loss: 0.068874791264534
step: 100, loss: 0.02269073761999607
step: 110, loss: 0.0019668200984597206
step: 120, loss: 0.045550089329481125
step: 130, loss: 0.006712373346090317
step: 140, loss: 0.000474466331070289
step: 150, loss: 0.018300363793969154
step: 160, loss: 0.0010436190059408545
step: 170, loss: 0.045392900705337524
step: 180, loss: 0.0009177344036288559
step: 190, loss: 0.001065362710505724
step: 200, loss: 0.002045310102403164
step: 210, loss: 0.006552683189511299

```

```

step: 220, loss: 0.0012736173812299967
step: 230, loss: 0.006179110147058964
step: 240, loss: 0.0006477690185420215
step: 250, loss: 0.0006247099954634905
step: 260, loss: 0.003831573761999607
step: 270, loss: 0.0009733483893796802
step: 280, loss: 0.002024790970608592
step: 290, loss: 0.1651114523410797
step: 300, loss: 0.0055348067544440069
step: 310, loss: 0.008496585302054882
step: 320, loss: 0.12145639210939407
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 330, loss: 0.037951063364744186
step: 340, loss: 0.016631968319416046
step: 350, loss: 0.004856766201555729
step: 360, loss: 0.0013891402631998062
step: 370, loss: 0.016128098592162132
step: 380, loss: 0.0020567590836435556
step: 390, loss: 0.00030411151237785816
step: 400, loss: 0.020665867254137993
step: 410, loss: 0.0034339562989771366
step: 420, loss: 0.002553262049332261
step: 430, loss: 0.008420160040259361
step: 440, loss: 0.013160738162696362
step: 450, loss: 0.04649581387639046
step: 460, loss: 0.00048246842925436795
step: 470, loss: 0.002129775006324053
step: 480, loss: 0.0030854654032737017
step: 490, loss: 0.0004895806778222322
step: 500, loss: 0.004125566687434912
step: 510, loss: 0.048987697809934616
step: 520, loss: 0.030894720926880836
step: 530, loss: 0.006270528305321932
/usr/local/lib/python3.7/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=0.944550669216061, f1=0.927038626609442,
best_f1=0.9439555349698935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013078608317300677
step: 10, loss: 0.0008774392190389335
step: 20, loss: 0.06775864213705063
step: 30, loss: 0.0042925006709992886
step: 40, loss: 0.0762103870511055
step: 50, loss: 0.03617304190993309
step: 60, loss: 0.0006635812460444868

```

step: 70, loss: 0.0050073652528226376
step: 80, loss: 0.010903864167630672
step: 90, loss: 0.00902643147855997
step: 100, loss: 0.0032286704517900944
step: 110, loss: 0.0005723875947296619
step: 120, loss: 0.003026435384526849
step: 130, loss: 0.0019277977989986539
step: 140, loss: 0.002599510829895735
step: 150, loss: 0.00039234518771991134
step: 160, loss: 0.0024160321336239576
step: 170, loss: 0.001912617706693709
step: 180, loss: 0.008116920478641987
step: 190, loss: 0.0020753296557813883
step: 200, loss: 0.02294228784739971
step: 210, loss: 0.04057176783680916
step: 220, loss: 0.00203108973801136
step: 230, loss: 0.007131707388907671
step: 240, loss: 0.05180423706769943
step: 250, loss: 0.021984893828630447
step: 260, loss: 0.0013233214849606156
step: 270, loss: 0.0709291473031044
step: 280, loss: 0.08680293709039688
step: 290, loss: 0.0007456646417267621
step: 300, loss: 0.0013159452937543392
step: 310, loss: 0.002757434267550707
step: 320, loss: 0.0039162649773061275
step: 330, loss: 0.0006965687498450279
step: 340, loss: 0.1561986207962036
step: 350, loss: 0.05404125526547432
step: 360, loss: 0.03802133724093437
step: 370, loss: 0.004601365886628628
step: 380, loss: 0.0006424042512662709
step: 390, loss: 0.013910377398133278
step: 400, loss: 0.006887323223054409
step: 410, loss: 0.0004804292693734169
step: 420, loss: 0.03565295785665512
step: 430, loss: 0.0018754296470433474
step: 440, loss: 0.0017535239458084106
step: 450, loss: 0.0007672401843592525
step: 460, loss: 0.00016873482672963291
step: 470, loss: 0.0012478386051952839
step: 480, loss: 0.009396448731422424
step: 490, loss: 0.02759546972811222
step: 500, loss: 0.038452524691820145
step: 510, loss: 0.004244944080710411
step: 520, loss: 0.011212628334760666
step: 530, loss: 0.006144207902252674
/usr/local/lib/python3.7/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=0.947077772664519, f1=0.934622467771639,
best_f1=0.9439555349698935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002795423788484186
step: 10, loss: 0.0001975266932277009
step: 20, loss: 0.000684777507558465
step: 30, loss: 0.0010727188782766461
step: 40, loss: 0.030261972919106483
step: 50, loss: 0.010244976729154587
step: 60, loss: 0.0005870354361832142
step: 70, loss: 0.008611110039055347
step: 80, loss: 0.004362436011433601
step: 90, loss: 0.0005308038671500981
step: 100, loss: 4.616596197593026e-05
step: 110, loss: 0.001203848165459931
step: 120, loss: 0.00029556165100075305
step: 130, loss: 0.00919285137206316
step: 140, loss: 5.362946103559807e-05
step: 150, loss: 0.0003632190346252173
step: 160, loss: 0.0026235352270305157
step: 170, loss: 0.0006043165922164917
step: 180, loss: 0.0042946008034050465
step: 190, loss: 0.0012049662182107568
step: 200, loss: 0.0013670535990968347
step: 210, loss: 0.0009985992219299078
step: 220, loss: 0.00014822991215623915
step: 230, loss: 0.02259596809744835
step: 240, loss: 0.002889778232201934
step: 250, loss: 0.005017393734306097
step: 260, loss: 0.00181643757969141
step: 270, loss: 0.0021696470212191343
step: 280, loss: 0.0027308182325214148
step: 290, loss: 0.00035685987677425146
step: 300, loss: 0.006150365341454744
step: 310, loss: 0.0003499095037113875
step: 320, loss: 0.00214031501673162
step: 330, loss: 0.008439745754003525
step: 340, loss: 0.014311681501567364
step: 350, loss: 0.005326885264366865
step: 360, loss: 0.014057882130146027
step: 370, loss: 0.001798818469978869
step: 380, loss: 0.002243570052087307
step: 390, loss: 0.0006820968119427562
step: 400, loss: 0.007949551567435265

```

```

step: 410, loss: 0.0022618232760578394
step: 420, loss: 0.0007869929540902376
step: 430, loss: 0.00012357735249679536
step: 440, loss: 0.1155693382024765
step: 450, loss: 0.0035247341729700565
step: 460, loss: 0.0007296453113667667
step: 470, loss: 0.01189132034778595
step: 480, loss: 0.034822091460227966
step: 490, loss: 0.0009410661295987666
step: 500, loss: 0.007238848600536585
step: 510, loss: 0.04706745594739914
step: 520, loss: 0.007702121511101723
step: 530, loss: 0.059655237942934036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9482439926062847, f1=0.9367789570835257,
best_f1=0.9367789570835257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008139782003127038
step: 10, loss: 0.017889460548758507
step: 20, loss: 0.00049936881987378
step: 30, loss: 0.032479673624038696
step: 40, loss: 0.0012040002038702369
step: 50, loss: 0.010626065544784069
step: 60, loss: 0.0019219653913751245
step: 70, loss: 0.00017805688548833132
step: 80, loss: 0.006365919020026922
step: 90, loss: 0.0011649479856714606
step: 100, loss: 0.000238623411860317
step: 110, loss: 0.00023493063054047525
step: 120, loss: 0.0012607518583536148
step: 130, loss: 0.037086278200149536
step: 140, loss: 0.0012610412668436766
step: 150, loss: 0.026543473824858665
step: 160, loss: 0.004255908541381359
step: 170, loss: 0.0007907228427939117
step: 180, loss: 0.0001001711716526188
step: 190, loss: 0.0024075335822999477
step: 200, loss: 0.0010803574696183205
step: 210, loss: 0.01603306084871292
step: 220, loss: 0.00128068495541811

```

```

step: 230, loss: 0.00023617841361556202
step: 240, loss: 0.0025841612368822098
step: 250, loss: 0.004370700102299452
step: 260, loss: 0.0002686180523596704
step: 270, loss: 0.0015906778862699866
step: 280, loss: 0.030982818454504013
step: 290, loss: 0.0007213752251118422
step: 300, loss: 0.0021554846316576004
step: 310, loss: 0.002059229416772723
step: 320, loss: 0.033265236765146255
step: 330, loss: 0.0010270723141729832
step: 340, loss: 0.0008808872662484646
step: 350, loss: 0.0006196614704094827
step: 360, loss: 0.0024703899398446083
step: 370, loss: 0.0005856921197846532
step: 380, loss: 0.0006812146166339517
step: 390, loss: 0.11741770803928375
step: 400, loss: 0.0003220649668946862
step: 410, loss: 0.0004735249967779964
step: 420, loss: 0.0032224440947175026
step: 430, loss: 0.018923815339803696
step: 440, loss: 0.011284041218459606
step: 450, loss: 0.0012730739545077085
step: 460, loss: 0.0009246696718037128
step: 470, loss: 0.00098162773065269
step: 480, loss: 8.979669655673206e-05
step: 490, loss: 0.10230732709169388
step: 500, loss: 0.01689630001783371
step: 510, loss: 0.023352500051259995
step: 520, loss: 0.0025941035710275173
step: 530, loss: 0.01772943139076233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9505597014925373, f1=0.9402985074626865,
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030193100683391094
step: 10, loss: 0.0004056826001033187
step: 20, loss: 0.00439093355089426
step: 30, loss: 0.0130064832046628
step: 40, loss: 0.004701430909335613

```

step: 50, loss: 0.00012064316979376599
step: 60, loss: 0.0001558809308335185
step: 70, loss: 0.005428020376712084
step: 80, loss: 0.0036008923780173063
step: 90, loss: 3.74379160348326e-05
step: 100, loss: 0.0003888345672748983
step: 110, loss: 0.0005379898939281702
step: 120, loss: 0.0016236932715401053
step: 130, loss: 0.022056259214878082
step: 140, loss: 0.0014591979561373591
step: 150, loss: 0.0006054423865862191
step: 160, loss: 0.0010221782140433788
step: 170, loss: 0.006541276816278696
step: 180, loss: 0.0007065017125569284
step: 190, loss: 0.0005621652817353606
step: 200, loss: 0.005750495009124279
step: 210, loss: 0.0026747360825538635
step: 220, loss: 0.00041831054841168225
step: 230, loss: 0.00036395119968801737
step: 240, loss: 0.00018193494179286063
step: 250, loss: 0.0025442102923989296
step: 260, loss: 0.010549026541411877
step: 270, loss: 5.907580634811893e-05
step: 280, loss: 0.00027617282466962934
step: 290, loss: 0.02650558203458786
step: 300, loss: 0.0003853222879115492
step: 310, loss: 0.010382300242781639
step: 320, loss: 0.0029251333326101303
step: 330, loss: 0.0027623495552688837
step: 340, loss: 0.00048511798377148807
step: 350, loss: 0.014466504566371441
step: 360, loss: 0.0005875502829439938
step: 370, loss: 0.0006120207253843546
step: 380, loss: 7.761686720186844e-05
step: 390, loss: 0.00033288446138612926
step: 400, loss: 0.005880423821508884
step: 410, loss: 0.00012707577843684703
step: 420, loss: 0.00011557240213733166
step: 430, loss: 0.0006837662076577544
step: 440, loss: 0.008833403699100018
step: 450, loss: 0.0006761419354006648
step: 460, loss: 0.00037078335299156606
step: 470, loss: 7.802057371009141e-05
step: 480, loss: 9.129514364758506e-05
step: 490, loss: 0.03192351385951042
step: 500, loss: 0.0012609506957232952
step: 510, loss: 0.0016084861708804965
step: 520, loss: 0.005223341751843691


```

step: 530, loss: 0.0005172122619114816
/usr/local/lib/python3.7/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=0.9410125406409661, f1=0.9430291801760075,
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011151331476867199
step: 10, loss: 0.0002601266896817833
step: 20, loss: 8.399529178859666e-05
step: 30, loss: 0.00024133548140525818
step: 40, loss: 0.0004950780421495438
step: 50, loss: 0.0004546429554466158
step: 60, loss: 0.00012437016994226724
step: 70, loss: 0.0001107377975131385
step: 80, loss: 0.00019385185441933572
step: 90, loss: 0.0005286990781314671
step: 100, loss: 0.0004346010973677039
step: 110, loss: 5.984608651488088e-05
step: 120, loss: 0.0010603107511997223
step: 130, loss: 0.0010047588730230927
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 140, loss: 0.012851848267018795
step: 150, loss: 0.0008160461438819766
step: 160, loss: 0.0031727557070553303
step: 170, loss: 0.0036447823513299227
step: 180, loss: 0.005628697108477354
step: 190, loss: 0.0009553141426295042
step: 200, loss: 0.0007246664026752114
step: 210, loss: 0.0008706926601007581
step: 220, loss: 0.02074907347559929
step: 230, loss: 0.0015237218467518687
step: 240, loss: 0.00013049854896962643
step: 250, loss: 0.000156427311594598
step: 260, loss: 0.001643929979763925
step: 270, loss: 0.0012457428965717554
step: 280, loss: 0.10191556811332703
step: 290, loss: 0.0005060902913101017
step: 300, loss: 0.00015209565754048526
step: 310, loss: 0.0008460783283226192
step: 320, loss: 0.0008181790472008288
step: 330, loss: 0.0009483667672611773
step: 340, loss: 0.029028335586190224
step: 350, loss: 0.0006725272396579385
step: 360, loss: 0.00013368792133405805
step: 370, loss: 0.001003874815069139

```

```

step: 380, loss: 0.0005110168131068349
step: 390, loss: 0.0006240950315259397
step: 400, loss: 0.0016069056000560522
step: 410, loss: 0.0018682136433199048
step: 420, loss: 0.00016230408800765872
step: 430, loss: 0.00013037899043411016
step: 440, loss: 0.00624297559261322
step: 450, loss: 6.113530253060162e-05
step: 460, loss: 0.00021506519988179207
step: 470, loss: 7.506393012590706e-05
step: 480, loss: 7.478124462068081e-05
step: 490, loss: 0.00015154748689383268
step: 500, loss: 0.009390168823301792
step: 510, loss: 5.518314355867915e-05
step: 520, loss: 0.00011018406075891107
step: 530, loss: 0.0010306520853191614
/usr/local/lib/python3.7/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=0.9421487603305785, f1=0.9413382218148486,
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.218915298930369e-05
step: 10, loss: 0.028453830629587173
step: 20, loss: 0.00025943017681129277
step: 30, loss: 0.0016716246027499437
step: 40, loss: 0.001500359969213605
step: 50, loss: 0.010878799483180046
step: 60, loss: 0.026540609076619148
step: 70, loss: 5.024405618314631e-05
step: 80, loss: 0.0009628729894757271
step: 90, loss: 3.8791069528087974e-05
step: 100, loss: 0.004561766516417265
step: 110, loss: 0.0005187960341572762
step: 120, loss: 0.00016857354785315692
step: 130, loss: 0.005318029783666134
step: 140, loss: 0.001107497955672443
step: 150, loss: 9.265524568036199e-05
step: 160, loss: 8.992934453999624e-05
step: 170, loss: 0.0013785588089376688
step: 180, loss: 0.0002472220512572676
step: 190, loss: 0.000692379311658442
step: 200, loss: 0.018539994955062866
step: 210, loss: 0.00011610710498644039
step: 220, loss: 0.0033344461116939783
step: 230, loss: 0.00028673477936536074

```

```

step: 240, loss: 0.00046502857003360987
step: 250, loss: 4.959207581123337e-05
step: 260, loss: 5.8676512708188966e-05
step: 270, loss: 0.002044510096311569
step: 280, loss: 0.0013055027229711413
step: 290, loss: 0.0021089529618620872
step: 300, loss: 0.002641493221744895
step: 310, loss: 0.00024145559291355312
step: 320, loss: 0.0028619880322366953
step: 330, loss: 0.00035820313496515155
step: 340, loss: 0.014051929116249084
step: 350, loss: 0.0006365827866829932
step: 360, loss: 0.0006963726482354105
step: 370, loss: 0.0013735020766034722
step: 380, loss: 0.00015450951468665153
step: 390, loss: 0.000453173037385568
step: 400, loss: 0.0005474817007780075
step: 410, loss: 0.0009758748346939683
step: 420, loss: 0.0004737314011435956
step: 430, loss: 0.002254580380395055
step: 440, loss: 0.0037001422606408596
step: 450, loss: 0.0002594192628748715
step: 460, loss: 0.0015004019951447845
step: 470, loss: 9.48207889450714e-05
step: 480, loss: 0.0009276974014937878
step: 490, loss: 0.0003009536594618112
step: 500, loss: 0.0008321081404574215
step: 510, loss: 0.001538340002298355
step: 520, loss: 3.5670269426191226e-05
step: 530, loss: 0.00027888055774383247
/usr/local/lib/python3.7/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=0.943239501615136, f1=0.9379310344827586,
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033127001370303333
step: 10, loss: 3.4245269489474595e-05
step: 20, loss: 0.0011116870446130633
step: 30, loss: 0.004181908443570137
step: 40, loss: 0.00016270895139314234
step: 50, loss: 0.0886593759059906
step: 60, loss: 0.0026187037583440542
step: 70, loss: 0.000768741883803159
step: 80, loss: 0.0007192462799139321
step: 90, loss: 0.010450397618114948

```

```
step: 100, loss: 0.001587169710546732
step: 110, loss: 0.005227155052125454
step: 120, loss: 6.126450898591429e-05
step: 130, loss: 0.00018085197370965034
step: 140, loss: 0.0014068171149119735
step: 150, loss: 0.00022167229326441884
step: 160, loss: 7.724702300038189e-05
step: 170, loss: 0.00015331583563238382
step: 180, loss: 0.00025086922687478364
step: 190, loss: 0.0005785523680970073
step: 200, loss: 0.00014867258141748607
step: 210, loss: 0.00022889007232151926
step: 220, loss: 3.1261268304660916e-05
step: 230, loss: 7.715155516052619e-05
step: 240, loss: 0.00021926709450781345
step: 250, loss: 0.00011595517571549863
step: 260, loss: 5.0355385610600933e-05
step: 270, loss: 5.7347755500813946e-05
step: 280, loss: 0.002104220213368535
step: 290, loss: 0.0029597890097647905
step: 300, loss: 0.00036139957956038415
step: 310, loss: 0.00011949393228860572
step: 320, loss: 0.0009443858289159834
step: 330, loss: 0.0003383801958989352
step: 340, loss: 7.838482270017266e-05
step: 350, loss: 0.00011369449930498376
step: 360, loss: 0.024678558111190796
step: 370, loss: 0.00016857919399626553
step: 380, loss: 5.7122069847537205e-05
step: 390, loss: 0.009059443138539791
step: 400, loss: 2.4638631657580845e-05
step: 410, loss: 0.0012737573124468327
step: 420, loss: 0.015434378758072853
step: 430, loss: 0.022894591093063354
step: 440, loss: 0.0032198706176131964
step: 450, loss: 0.0006112062837928534
step: 460, loss: 0.000824771064799279
step: 470, loss: 0.0049249110743403435
step: 480, loss: 0.0017389446729794145
step: 490, loss: 0.001456206082366407
step: 500, loss: 0.003914260771125555
step: 510, loss: 0.00326289189979434
step: 520, loss: 8.59415958984755e-05
step: 530, loss: 0.23017756640911102
/usr/local/lib/python3.7/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=0.9476584022038568, f1=0.9424657534246577,
```

```
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015718918293714523
step: 10, loss: 0.00021901940635871142
step: 20, loss: 0.004183280747383833
step: 30, loss: 0.0004955038311891258
step: 40, loss: 6.295557250268757e-05
step: 50, loss: 0.000172117113834247
step: 60, loss: 0.013886659406125546
step: 70, loss: 0.0002716902527026832
step: 80, loss: 0.0010038106702268124
step: 90, loss: 0.013988396152853966
step: 100, loss: 0.007311410270631313
step: 110, loss: 0.00016279277042485774
step: 120, loss: 0.003270109184086323
step: 130, loss: 0.00012845598394051194
step: 140, loss: 5.7654702686704695e-05
step: 150, loss: 0.0014918290544301271
step: 160, loss: 3.0028490073163994e-05
step: 170, loss: 0.00424310052767396
step: 180, loss: 0.0009806439047679305
step: 190, loss: 0.0003072282124776393
step: 200, loss: 9.822621359489858e-05
step: 210, loss: 0.00018813807400874794
step: 220, loss: 7.676428504055366e-05
step: 230, loss: 0.005205960478633642
step: 240, loss: 0.00021708215354010463
step: 250, loss: 0.00020759162725880742
step: 260, loss: 0.00015310331946238875
step: 270, loss: 0.00010001331975217909
step: 280, loss: 0.0005002537509426475
step: 290, loss: 0.027032289654016495
step: 300, loss: 0.00010161507088923827
step: 310, loss: 0.0008983528823591769
step: 320, loss: 0.004215933382511139
step: 330, loss: 0.0019288365729153156
step: 340, loss: 0.0015114230336621404
step: 350, loss: 0.0004009544791188091
step: 360, loss: 2.9905444534961134e-05
step: 370, loss: 0.005279839970171452
step: 380, loss: 0.00013056147145107388
step: 390, loss: 0.001632555155083537
step: 400, loss: 0.0010531748412176967
step: 410, loss: 0.002923036925494671
step: 420, loss: 5.6170996685978025e-05
step: 430, loss: 0.0165165513753891
```

```

step: 440, loss: 0.00012059631262673065
step: 450, loss: 0.0005768531118519604
step: 460, loss: 0.0017606235342100263
step: 470, loss: 0.00012940082524437457
step: 480, loss: 0.0019013747805729508
step: 490, loss: 0.0010401408653706312
step: 500, loss: 0.0003859555290546268
step: 510, loss: 0.0010723811574280262
step: 520, loss: 0.0016155792400240898
step: 530, loss: 0.00022299772535916418
/usr/local/lib/python3.7/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=0.9486348912540491, f1=0.9422632794457274,
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013464185176417232
step: 10, loss: 4.018565232399851e-05
step: 20, loss: 0.003085819073021412
step: 30, loss: 4.3579268094617873e-05
step: 40, loss: 0.00015357301163021475
step: 50, loss: 0.0037952482234686613
step: 60, loss: 2.7923531888518482e-05
step: 70, loss: 0.001314960652962327
step: 80, loss: 5.141804649611004e-05
step: 90, loss: 0.0005231865798123181
step: 100, loss: 0.0002039415849139914
step: 110, loss: 0.00011824697867268696
step: 120, loss: 0.00289644836448133
step: 130, loss: 0.010952156037092209
step: 140, loss: 0.0016621516551822424
step: 150, loss: 6.229297287063673e-05
step: 160, loss: 0.0001524297404102981
step: 170, loss: 0.0010639779502525926
step: 180, loss: 4.142095349379815e-05
step: 190, loss: 9.774582576937973e-05
step: 200, loss: 0.000201451126486063
step: 210, loss: 3.764961729757488e-05
step: 220, loss: 0.00020321564807090908
step: 230, loss: 0.0066797430627048016
step: 240, loss: 0.00013487094838637859
step: 250, loss: 0.0023418704513460398
step: 260, loss: 0.0021837488748133183
step: 270, loss: 0.0009161950438283384
step: 280, loss: 0.00029460960649885237
step: 290, loss: 0.0009611837449483573

```

```

step: 300, loss: 0.0012230115244165063
step: 310, loss: 0.10992524027824402
step: 320, loss: 0.0019190995953977108
step: 330, loss: 0.0011069212341681123
step: 340, loss: 9.314336784882471e-05
step: 350, loss: 0.0011065109865739942
step: 360, loss: 0.013202709145843983
step: 370, loss: 5.101356509840116e-05
step: 380, loss: 0.00016277635586448014
step: 390, loss: 0.05128394812345505
step: 400, loss: 0.001019367715343833
step: 410, loss: 0.00018400604312773794
step: 420, loss: 0.0007771337986923754
step: 430, loss: 3.617669062805362e-05
step: 440, loss: 0.0019375178962945938
step: 450, loss: 0.00015573165728710592
step: 460, loss: 0.0003297730290796608
step: 470, loss: 0.0005538191180676222
step: 480, loss: 0.0025557836052030325
step: 490, loss: 2.4589975510025397e-05
step: 500, loss: 0.0006348148453980684
step: 510, loss: 0.04378880187869072
step: 520, loss: 0.00024238364130724221
step: 530, loss: 0.0031421915628015995
/usr/local/lib/python3.7/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=0.948087431693989, f1=0.9394760614272809,
best_f1=0.9402985074626865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.5830480808508582e-05
step: 10, loss: 3.867108898703009e-05
step: 20, loss: 0.00011218279541935772
step: 30, loss: 0.00047990932944230735
step: 40, loss: 0.07597818970680237
step: 50, loss: 0.0007449730765074492
step: 60, loss: 4.5876658987253904e-05
step: 70, loss: 0.00040733328205533326
step: 80, loss: 0.0005504100699909031
step: 90, loss: 9.872250666376203e-05
step: 100, loss: 0.0002124317252309993
step: 110, loss: 0.0006108639063313603
step: 120, loss: 0.003963451832532883
step: 130, loss: 0.06811362504959106
step: 140, loss: 3.7277251976775005e-05
step: 150, loss: 0.00034302700078114867

```

```

step: 160, loss: 8.350797725142911e-05
step: 170, loss: 0.004256750922650099
step: 180, loss: 0.0028855171985924244
step: 190, loss: 7.162239489844069e-05
step: 200, loss: 5.9596815844997764e-05
step: 210, loss: 0.00048691959818825126
step: 220, loss: 5.4969594202702865e-05
step: 230, loss: 0.008095129393041134
step: 240, loss: 0.00011221024760743603
step: 250, loss: 4.198082388029434e-05
step: 260, loss: 0.0014007112476974726
step: 270, loss: 5.5741787946317345e-05
step: 280, loss: 4.0979080949909985e-05
step: 290, loss: 1.8100956367561594e-05
step: 300, loss: 0.00020812494040001184
step: 310, loss: 0.000194115360500291
step: 320, loss: 0.0002751545107457787
step: 330, loss: 0.00044100871309638023
step: 340, loss: 8.668829104863107e-05
step: 350, loss: 2.6634808818926103e-05
step: 360, loss: 0.0030076326802372932
step: 370, loss: 0.0002291148848598823
step: 380, loss: 0.00017854692123364657
step: 390, loss: 0.00027443180442787707
step: 400, loss: 0.0001610801846254617
step: 410, loss: 0.0004252489598002285
step: 420, loss: 0.00014213263057172298
step: 430, loss: 0.00011571541836019605
step: 440, loss: 0.01089342962950468
step: 450, loss: 2.1580277461907826e-05
step: 460, loss: 0.00013528933050110936
step: 470, loss: 0.0013136824127286673
step: 480, loss: 5.972296276013367e-05
step: 490, loss: 2.3836859327275306e-05
step: 500, loss: 0.00017698090232443064
step: 510, loss: 0.000587764719966799
step: 520, loss: 2.3960534235811792e-05
step: 530, loss: 0.0001153593257186003
/usr/local/lib/python3.7/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=0.9484820607175714, f1=0.9424131627056673,
best_f1=0.9402985074626865
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:


```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 235.24it/s]
load_f1 = 0.9499766245909304
real_f1 = 0.9485500467726847
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, 235.56it/s]

```

2 BASELINE

2.1 BASELINE STRUCTURED

2.1.1 Beer - Running the matcher

```

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

```

```

--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm bert-base-uncased \
--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

An input tensor was not cuda.

```

    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5493285655975342
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3703703703703704, f1=0.41025641025641024,
best_f1=0.41025641025641024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.47930142283439636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5263157894736842, f1=0.3934426229508196,
best_f1=0.3934426229508196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4375120997428894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5641025641025641, f1=0.39999999999999997,
best_f1=0.39999999999999997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3387130796909332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.588235294117647, f1=0.5116279069767441,
best_f1=0.5116279069767441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22148816287517548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6875000000000001, f1=0.5945945945945946,
best_f1=0.5945945945945946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18645012378692627
/usr/local/lib/python3.7/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=0.6086956521739131, f1=0.7200000000000001,
best_f1=0.5945945945945946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15051722526550293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.7741935483870968, f1=0.6666666666666667,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022493980824947357
/usr/local/lib/python3.7/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=0.5454545454545454, f1=0.7200000000000001,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025108305271714926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7857142857142857, f1=0.7058823529411764,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1207951009273529
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.6470588235294117,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005484204273670912
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.6857142857142857,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015420910902321339
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7999999999999999, f1=0.6666666666666667,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006273516453802586
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6666666666666667,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005531961563974619
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6666666666666667,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008569241501390934
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6666666666666667,
best_f1=0.6857142857142857
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, 139046.14it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

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/

```

Downloading: 100% 433/433 [00:00<00:00, 394kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 360kB/s]
Downloading: 100% 440M/440M [00:08<00:00, 49.4MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

step: 0, loss: 0.6323108077049255
step: 10, loss: 0.6318007707595825
step: 20, loss: 0.3705366253852844
step: 30, loss: 0.1814804971218109
step: 40, loss: 0.1457262933254242
step: 50, loss: 0.025757424533367157
step: 60, loss: 0.033064454793930054
step: 70, loss: 0.03568490222096443

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.14440089464187622


```

step: 90, loss: 0.17474599182605743
step: 100, loss: 0.007030672859400511
step: 110, loss: 0.210466668009758
step: 120, loss: 0.011536875739693642
step: 130, loss: 0.005657666828483343
step: 140, loss: 0.003935035318136215
step: 150, loss: 0.007173934951424599
step: 160, loss: 0.05143074318766594
step: 170, loss: 0.1309310644865036
step: 180, loss: 0.03851013258099556
step: 190, loss: 0.011646075174212456
step: 200, loss: 0.134754940867424
step: 210, loss: 0.0032973592169582844
step: 220, loss: 0.0057314918376505375
step: 230, loss: 0.0071966517716646194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841628959276018, f1=0.9829738933030647,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017588584451004863
step: 10, loss: 0.021488569676876068
step: 20, loss: 0.134184792637825
step: 30, loss: 0.17799067497253418
step: 40, loss: 0.09229615330696106
step: 50, loss: 0.011168969795107841
step: 60, loss: 0.003647582372650504
step: 70, loss: 0.0782218649983406
step: 80, loss: 0.017864780500531197
step: 90, loss: 0.003031366039067507
step: 100, loss: 0.004582575522363186
step: 110, loss: 0.12011457234621048
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.00040419475408270955
step: 130, loss: 0.00227834889665246
step: 140, loss: 0.0015656378818675876
step: 150, loss: 0.00476599857211113
step: 160, loss: 0.0061200945638120174
step: 170, loss: 0.0015795885119587183
step: 180, loss: 0.08069151639938354
step: 190, loss: 0.08313318341970444

```

```

step: 200, loss: 0.006372259464114904
step: 210, loss: 0.0013681984273716807
step: 220, loss: 0.04462727904319763
step: 230, loss: 0.006205535028129816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9887387387387387, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002398143755272031
step: 10, loss: 0.0024643330834805965
step: 20, loss: 0.003036231268197298
step: 30, loss: 0.08119514584541321
step: 40, loss: 0.15170975029468536
step: 50, loss: 0.005599288269877434
step: 60, loss: 0.002723380923271179
step: 70, loss: 0.0012343612033873796
step: 80, loss: 0.00040756785892881453
step: 90, loss: 0.03900335356593132
step: 100, loss: 0.0004921568906866014
step: 110, loss: 0.003751930082216859
step: 120, loss: 0.01040804386138916
step: 130, loss: 0.00037110704579390585
step: 140, loss: 0.001190214417874813
step: 150, loss: 0.001324610086157918
step: 160, loss: 0.002822434762492776
step: 170, loss: 0.00840711873024702
step: 180, loss: 0.028852490708231926
step: 190, loss: 0.0027312415186315775
step: 200, loss: 0.00548789044842124
step: 210, loss: 0.0017006794223561883
step: 220, loss: 0.000566553499083966
step: 230, loss: 0.00043291517067700624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921436588103255, f1=0.9865470852017937,

```

```

best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00034498650347813964
step: 10, loss: 0.00034788623452186584
step: 20, loss: 0.0003632169682532549
step: 30, loss: 0.0006650419090874493
step: 40, loss: 0.003352970350533724
step: 50, loss: 0.00025919597828760743
step: 60, loss: 0.00038653070805594325
step: 70, loss: 0.0003277936193626374
step: 80, loss: 0.026498982682824135
step: 90, loss: 0.0013113352470099926
step: 100, loss: 0.00031642010435462
step: 110, loss: 0.000561875116545707
step: 120, loss: 0.06291807442903519
step: 130, loss: 0.0007969868602231145
step: 140, loss: 0.0023518051020801067
step: 150, loss: 0.07349002361297607
step: 160, loss: 0.002375139156356454
step: 170, loss: 0.02786041609942913
step: 180, loss: 0.0005144757451489568
step: 190, loss: 0.005785717628896236
step: 200, loss: 0.0005149073549546301
step: 210, loss: 0.05586298927664757
step: 220, loss: 0.016717027872800827
step: 230, loss: 0.034148119390010834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9865168539325843, f1=0.979591836734694,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005030049476772547
step: 10, loss: 0.024727223441004753
step: 20, loss: 0.04569259658455849
step: 30, loss: 0.0004894825397059321
step: 40, loss: 0.01461218111217022
step: 50, loss: 0.00022172529133968055
step: 60, loss: 0.006872154306620359
step: 70, loss: 0.002351323375478387
step: 80, loss: 0.00019496653112582862
step: 90, loss: 0.002289149910211563
step: 100, loss: 0.00022502588399220258
step: 110, loss: 0.00013288618356455117

```

```

step: 120, loss: 9.911808592732996e-05
step: 130, loss: 0.0007463739602826536
step: 140, loss: 0.00043875479605048895
step: 150, loss: 0.0003665466792881489
step: 160, loss: 0.0014629310462623835
step: 170, loss: 0.0071861534379422665
step: 180, loss: 0.000708176929038018
step: 190, loss: 0.0012931902892887592
step: 200, loss: 0.0032518624793738127
step: 210, loss: 0.00042619684245437384
step: 220, loss: 0.0007804164779372513
step: 230, loss: 0.00026441054069437087
/usr/local/lib/python3.7/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=0.9887640449438202,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008692018454894423
step: 10, loss: 0.0008259998285211623
step: 20, loss: 0.001763767097145319
step: 30, loss: 0.00034805884934030473
step: 40, loss: 0.0009263046667911112
step: 50, loss: 0.00023198428971227258
step: 60, loss: 0.0001648802572162822
step: 70, loss: 0.00024910373031161726
step: 80, loss: 0.001468566944822669
step: 90, loss: 0.0005033523193560541
step: 100, loss: 0.029769619926810265
step: 110, loss: 0.001739682862535119
step: 120, loss: 0.0005133732338435948
step: 130, loss: 0.0018062799936160445
step: 140, loss: 0.00012883210729341954
step: 150, loss: 0.004170480649918318
step: 160, loss: 0.00030426454031839967
step: 170, loss: 0.00018677100888453424
step: 180, loss: 0.0295083187520504
step: 190, loss: 0.001492014736868441
step: 200, loss: 0.0002922893618233502
step: 210, loss: 0.00048156900447793305
step: 220, loss: 0.00015184156654868275
step: 230, loss: 0.06455713510513306
/usr/local/lib/python3.7/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=0.9910313901345291, f1=0.9855072463768116,

```

```

best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003436678380239755
step: 10, loss: 0.00013353975373320282
step: 20, loss: 0.0006387712201103568
step: 30, loss: 0.00014635881234426051
step: 40, loss: 0.00018473480304237455
step: 50, loss: 0.0015791127225384116
step: 60, loss: 0.00013898743782192469
step: 70, loss: 0.023489603772759438
step: 80, loss: 0.00025371499941684306
step: 90, loss: 7.6708645792678e-05
step: 100, loss: 0.0001512105081928894
step: 110, loss: 0.00013756650150753558
step: 120, loss: 0.17822912335395813
step: 130, loss: 0.00016825305647216737
step: 140, loss: 0.0003726343566086143
step: 150, loss: 0.0052201044745743275
step: 160, loss: 0.06554825603961945
step: 170, loss: 0.002957398071885109
step: 180, loss: 0.001735014608129859
step: 190, loss: 0.00035796710290014744
step: 200, loss: 0.010642717592418194
step: 210, loss: 6.58018107060343e-05
step: 220, loss: 0.0002155935944756493
step: 230, loss: 0.037829749286174774
/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9864864864864865,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.863006612751633e-05
step: 10, loss: 0.0003776528174057603
step: 20, loss: 8.306876407004893e-05
step: 30, loss: 0.0017914085183292627
step: 40, loss: 0.0002536701795179397
step: 50, loss: 0.0007484082016162574
step: 60, loss: 9.617459727451205e-05
step: 70, loss: 0.0006986379157751799
step: 80, loss: 0.00014796850155107677
step: 90, loss: 0.000507485936395824
step: 100, loss: 0.0001380612375214696
step: 110, loss: 0.07323545217514038

```

```

step: 120, loss: 0.0038226766046136618
step: 130, loss: 0.00012087801587767899
step: 140, loss: 8.187060302589089e-05
step: 150, loss: 0.00022016221191734076
step: 160, loss: 0.0003289382148068398
step: 170, loss: 6.455415859818459e-05
step: 180, loss: 0.00018068242934532464
step: 190, loss: 0.0001170701434602961
step: 200, loss: 9.656289330450818e-05
step: 210, loss: 0.00012200645869597793
step: 220, loss: 0.00010649113391991705
step: 230, loss: 0.00020140389096923172
/usr/local/lib/python3.7/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=0.9921259842519685, f1=0.9876819708846584,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.553024286404252e-05
step: 10, loss: 0.00017659277364145964
step: 20, loss: 0.0012384597212076187
step: 30, loss: 9.230526484316215e-05
step: 40, loss: 0.0678090751171112
step: 50, loss: 8.08402182883583e-05
step: 60, loss: 0.0005172500386834145
step: 70, loss: 0.0019485186785459518
step: 80, loss: 0.00014088630268815905
step: 90, loss: 0.00037882933975197375
step: 100, loss: 0.0004019145853817463
step: 110, loss: 0.00010138285142602399
step: 120, loss: 6.264662079047412e-05
step: 130, loss: 8.204727055272087e-05
step: 140, loss: 0.00935354270040989
step: 150, loss: 0.00036364770494401455
step: 160, loss: 6.0180907894391567e-05
step: 170, loss: 0.0002096561511280015
step: 180, loss: 0.00035948483855463564
step: 190, loss: 6.0166075854795054e-05
step: 200, loss: 6.170119013404474e-05
step: 210, loss: 0.0006454944959841669
step: 220, loss: 0.00010732511145761237
step: 230, loss: 0.00018267222912982106
/usr/local/lib/python3.7/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=0.9910714285714286, f1=0.9811738648947952,

```

```

best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.460674194386229e-05
step: 10, loss: 0.00014741664926987141
step: 20, loss: 9.718913497636095e-05
step: 30, loss: 0.0012329673627391458
step: 40, loss: 0.0002952116192318499
step: 50, loss: 6.867740012239665e-05
step: 60, loss: 0.014699381776154041
step: 70, loss: 0.000333546195179224
step: 80, loss: 7.420303154503927e-05
step: 90, loss: 0.006820885930210352
step: 100, loss: 9.964275523088872e-05
step: 110, loss: 0.00010720239515649155
step: 120, loss: 0.00028801767621189356
step: 130, loss: 7.43149284971878e-05
step: 140, loss: 0.0393211655318737
step: 150, loss: 0.014735261909663677
step: 160, loss: 3.623092925408855e-05
step: 170, loss: 0.00015269886353053153
step: 180, loss: 8.632188109913841e-05
step: 190, loss: 0.004737492185086012
step: 200, loss: 5.7775865570874885e-05
step: 210, loss: 8.98202633834444e-05
step: 220, loss: 5.1957849791506305e-05
step: 230, loss: 0.0003835907846223563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921787709497207, f1=0.9811320754716982,
best_f1=0.9811320754716982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001834666181821376
step: 10, loss: 5.7929264585254714e-05
step: 20, loss: 7.998820365173742e-05
step: 30, loss: 0.01646416075527668
step: 40, loss: 0.0009171520941890776
step: 50, loss: 0.0009849518537521362
step: 60, loss: 0.00015956032439135015
step: 70, loss: 0.0003339194809086621

```

```

step: 80, loss: 4.7483597882092e-05
step: 90, loss: 0.0003325370198581368
step: 100, loss: 5.9729332861024886e-05
step: 110, loss: 0.0038444004021584988
step: 120, loss: 3.762035703402944e-05
step: 130, loss: 4.630603871191852e-05
step: 140, loss: 5.5328691814793274e-05
step: 150, loss: 0.02768113650381565
step: 160, loss: 4.768932922161184e-05
step: 170, loss: 0.02056114934384823
step: 180, loss: 4.6169057895895094e-05
step: 190, loss: 4.414600334712304e-05
step: 200, loss: 0.0003837058611679822
step: 210, loss: 3.65588566637598e-05
step: 220, loss: 4.9201200454263017e-05
step: 230, loss: 6.389242480508983e-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=0.9899216125419933, f1=0.9843400447427293,
best_f1=0.9811320754716982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.741637167171575e-05
step: 10, loss: 4.041747160954401e-05
step: 20, loss: 5.1063769205939025e-05
step: 30, loss: 4.7804471250856295e-05
step: 40, loss: 5.1913018978666514e-05
step: 50, loss: 0.00010726058098953217
step: 60, loss: 0.0001966481504496187
step: 70, loss: 3.946386277675629e-05
step: 80, loss: 4.683519000536762e-05
step: 90, loss: 4.375747812446207e-05
step: 100, loss: 6.921072781551629e-05
step: 110, loss: 0.00010767908679554239
step: 120, loss: 6.65357947582379e-05
step: 130, loss: 4.0276423533214256e-05
step: 140, loss: 0.023205075412988663
step: 150, loss: 4.9978989409282804e-05
step: 160, loss: 4.069342321599834e-05
step: 170, loss: 0.00028531128191389143
step: 180, loss: 0.04153873771429062
step: 190, loss: 0.00022588929277844727
step: 200, loss: 2.7346681235940196e-05
step: 210, loss: 3.734475467354059e-05
step: 220, loss: 4.3031323002651334e-05
step: 230, loss: 0.027984585613012314

```



```

/usr/local/lib/python3.7/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=0.9887892376681614, f1=0.984304932735426,
best_f1=0.9811320754716982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011935082147829235
step: 10, loss: 6.778571696486324e-05
step: 20, loss: 9.247497655451298e-05
step: 30, loss: 0.000158140086568892
step: 40, loss: 5.687598968506791e-05
step: 50, loss: 6.405647582141683e-05
step: 60, loss: 4.749109939439222e-05
step: 70, loss: 2.830380981322378e-05
step: 80, loss: 3.844318780465983e-05
step: 90, loss: 6.697858771076426e-05
step: 100, loss: 2.904879511334002e-05
step: 110, loss: 0.022326668724417686
step: 120, loss: 0.028648316860198975
step: 130, loss: 4.615806392394006e-05
step: 140, loss: 5.0293787353439257e-05
step: 150, loss: 4.6292032493511215e-05
step: 160, loss: 4.0795584936859086e-05
step: 170, loss: 3.609676059568301e-05
step: 180, loss: 4.276354957255535e-05
step: 190, loss: 3.1690102332504466e-05
step: 200, loss: 5.396669439505786e-05
step: 210, loss: 3.101404581684619e-05
step: 220, loss: 7.574741175631061e-05
step: 230, loss: 4.843496571993455e-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=0.9887892376681614, f1=0.9854423292273236,
best_f1=0.9811320754716982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.747104685520753e-05
step: 10, loss: 0.01461595669388771
step: 20, loss: 3.883067620336078e-05
step: 30, loss: 6.738406955264509e-05
step: 40, loss: 0.0105152428150177
step: 50, loss: 4.152774272370152e-05
step: 60, loss: 7.947939593577757e-05
step: 70, loss: 5.945867087575607e-05

```

```

step: 80, loss: 5.0206337618874386e-05
step: 90, loss: 3.953464693040587e-05
step: 100, loss: 5.961778515484184e-05
step: 110, loss: 0.00010182803816860542
step: 120, loss: 2.011991455219686e-05
step: 130, loss: 5.595257243840024e-05
step: 140, loss: 5.0663995352806523e-05
step: 150, loss: 3.1105195375857875e-05
step: 160, loss: 0.0009146960801444948
step: 170, loss: 2.9525406716857105e-05
step: 180, loss: 5.4111715144244954e-05
step: 190, loss: 5.56629711354617e-05
step: 200, loss: 2.887759546865709e-05
step: 210, loss: 6.454421964008361e-05
step: 220, loss: 5.752571814809926e-05
step: 230, loss: 4.7017467295518145e-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=0.9887892376681614, f1=0.9843400447427293,
best_f1=0.9811320754716982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7383781343814917e-05
step: 10, loss: 2.2101688955444843e-05
step: 20, loss: 3.984417344327085e-05
step: 30, loss: 7.91921847849153e-05
step: 40, loss: 5.4713098506908864e-05
step: 50, loss: 0.05113434046506882
step: 60, loss: 0.00019488537509460002
step: 70, loss: 3.4476412110961974e-05
step: 80, loss: 3.326918158563785e-05
step: 90, loss: 0.00026156831881962717
step: 100, loss: 3.0445931770373136e-05
step: 110, loss: 2.8363390811136924e-05
step: 120, loss: 6.096565266489051e-05
step: 130, loss: 0.018456676974892616
step: 140, loss: 2.777880035864655e-05
step: 150, loss: 7.175163773354143e-05
step: 160, loss: 3.256101263104938e-05
step: 170, loss: 2.528657751099672e-05
step: 180, loss: 6.76176932756789e-05
step: 190, loss: 0.0001423422945663333
step: 200, loss: 6.519888847833499e-05
step: 210, loss: 3.4617751225596294e-05
step: 220, loss: 5.968333425698802e-05
step: 230, loss: 3.0121753297862597e-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=0.9887892376681614, f1=0.9831649831649831,
best_f1=0.9811320754716982
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:14, 165.00it/s]
load_f1 = 0.9921787709497207
real_f1 = 0.9921787709497207
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]

```

2.1.3 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: 0.6220229864120483
step: 10, loss: 0.5204026103019714
step: 20, loss: 0.45123526453971863
step: 30, loss: 0.04552942141890526
step: 40, loss: 0.20690588653087616
step: 50, loss: 0.1946549117565155
step: 60, loss: 0.11834309995174408
step: 70, loss: 0.0597916878759861
step: 80, loss: 0.026732252910733223
step: 90, loss: 0.12168753892183304
step: 100, loss: 0.007759185042232275
step: 110, loss: 0.10831107944250107
step: 120, loss: 0.03383738175034523
step: 130, loss: 0.07672528922557831
step: 140, loss: 0.03208855912089348
step: 150, loss: 0.04950945824384689
step: 160, loss: 0.019210198894143105
step: 170, loss: 0.14653688669204712
step: 180, loss: 0.05915548652410507
step: 190, loss: 0.01227243896573782
step: 200, loss: 0.13989822566509247
step: 210, loss: 0.07503893971443176
step: 220, loss: 0.14473164081573486
step: 230, loss: 0.14339780807495117
step: 240, loss: 0.09423361718654633
step: 250, loss: 0.035826440900564194
step: 260, loss: 0.05232656002044678
step: 270, loss: 0.005160799250006676

```

```

step: 280, loss: 0.027257995679974556
step: 290, loss: 0.08112740516662598
step: 300, loss: 0.04442750662565231
step: 310, loss: 0.18356379866600037
step: 320, loss: 0.07001284509897232
step: 330, loss: 0.010283250361680984
step: 340, loss: 0.025730768218636513
step: 350, loss: 0.04306721314787865
step: 360, loss: 0.0462871789932251
step: 370, loss: 0.09999199956655502
step: 380, loss: 0.007140554487705231
step: 390, loss: 0.21509721875190735
step: 400, loss: 0.16523852944374084
step: 410, loss: 0.0312829464673996
step: 420, loss: 0.022501766681671143
step: 430, loss: 0.13709954917430878
step: 440, loss: 0.0171003807336092
step: 450, loss: 0.0028508903924375772
step: 460, loss: 0.0035277113784104586
step: 470, loss: 0.09465762227773666
step: 480, loss: 0.037275295704603195
step: 490, loss: 0.06193116679787636
step: 500, loss: 0.08286254853010178
step: 510, loss: 0.05224613845348358
step: 520, loss: 0.032522156834602356
step: 530, loss: 0.00837781559675932
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9412861136999067, f1=0.9428172942817293,
best_f1=0.9428172942817293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1645800620317459
step: 10, loss: 0.05294647440314293
step: 20, loss: 0.009072313085198402
step: 30, loss: 0.04104965180158615
step: 40, loss: 0.08712183684110641
step: 50, loss: 0.14498786628246307
step: 60, loss: 0.0020352578721940517
step: 70, loss: 0.060350451618433
step: 80, loss: 0.06109953299164772
step: 90, loss: 0.018437812104821205

```

```
step: 100, loss: 0.006132568698376417
step: 110, loss: 0.0068924142979085445
step: 120, loss: 0.026416027918457985
step: 130, loss: 0.02779904566705227
step: 140, loss: 0.008943973109126091
step: 150, loss: 0.029324037954211235
step: 160, loss: 0.010673850774765015
step: 170, loss: 0.036962393671274185
step: 180, loss: 0.020480433478951454
step: 190, loss: 0.06902725994586945
step: 200, loss: 0.014214831404387951
step: 210, loss: 0.03292975574731827
step: 220, loss: 0.08898419886827469
step: 230, loss: 0.006584141869097948
step: 240, loss: 0.08329077064990997
step: 250, loss: 0.01646478660404682
step: 260, loss: 0.0018919638823717833
step: 270, loss: 0.17526189982891083
step: 280, loss: 0.017052508890628815
step: 290, loss: 0.025220435112714767
step: 300, loss: 0.18280796706676483
step: 310, loss: 0.010171854868531227
step: 320, loss: 0.09638208150863647
step: 330, loss: 0.02836466021835804
step: 340, loss: 0.007586915045976639
step: 350, loss: 0.0009929093066602945
step: 360, loss: 0.045821987092494965
step: 370, loss: 0.1034807488322258
step: 380, loss: 0.019202379509806633
step: 390, loss: 0.07898155599832535
step: 400, loss: 0.050033967941999435
step: 410, loss: 0.014346652664244175
step: 420, loss: 0.13833104074001312
step: 430, loss: 0.01147752907127142
step: 440, loss: 0.08867994695901871
step: 450, loss: 0.007375150453299284
step: 460, loss: 0.03804357349872589
step: 470, loss: 0.040114302188158035
step: 480, loss: 0.2991026043891907
step: 490, loss: 0.010328838601708412
step: 500, loss: 0.12179167568683624
step: 510, loss: 0.010730838403105736
step: 520, loss: 0.0290476456284523
step: 530, loss: 0.003634697524830699
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9447795823665893, f1=0.9416666666666667,

best_f1=0.9416666666666667

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

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

step: 0, loss: 0.0530112199485302

step: 10, loss: 0.10136982798576355

step: 20, loss: 0.1274033486843109

step: 30, loss: 0.04113495722413063

step: 40, loss: 0.0804159939289093

step: 50, loss: 0.006795635912567377

step: 60, loss: 0.011340384371578693

step: 70, loss: 0.0030027900356799364

step: 80, loss: 0.0016646446892991662

step: 90, loss: 0.008505373261868954

step: 100, loss: 0.13229916989803314

step: 110, loss: 0.0426144041121006

step: 120, loss: 0.004265810362994671

step: 130, loss: 0.004996461793780327

step: 140, loss: 0.0264773890376091

step: 150, loss: 0.023463737219572067

step: 160, loss: 0.01874222420156002

step: 170, loss: 0.11431229114532471

step: 180, loss: 0.009894225746393204

step: 190, loss: 0.015010195784270763

step: 200, loss: 0.024361146613955498

step: 210, loss: 0.046191874891519547

step: 220, loss: 0.055668845772743225

step: 230, loss: 0.07378114759922028

step: 240, loss: 0.001821873476728797

step: 250, loss: 0.03642292693257332

step: 260, loss: 0.009214146994054317

step: 270, loss: 0.0007683390285819769

step: 280, loss: 0.04102976992726326

step: 290, loss: 0.003690082347020507

step: 300, loss: 0.05305575579404831

step: 310, loss: 0.013635985553264618

step: 320, loss: 0.009373211301863194

step: 330, loss: 0.004902948159724474

step: 340, loss: 0.0029261731542646885

step: 350, loss: 0.0032268089707940817

step: 360, loss: 0.007700548507273197

step: 370, loss: 0.014936174266040325

step: 380, loss: 0.00345216179266572

step: 390, loss: 0.013958010822534561


```

step: 400, loss: 0.07230770587921143
step: 410, loss: 0.0043820892460644245
step: 420, loss: 0.06797308474779129
step: 430, loss: 0.004378196317702532
step: 440, loss: 0.0101764015853405
step: 450, loss: 0.08042647689580917
step: 460, loss: 0.03675726801156998
step: 470, loss: 0.013982388190925121
step: 480, loss: 0.0015218071639537811
step: 490, loss: 0.0044736601412296295
step: 500, loss: 0.01321484800428152
step: 510, loss: 0.0951448306441307
step: 520, loss: 0.07775821536779404
step: 530, loss: 0.08720158785581589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9455560725919032, f1=0.939435968562182,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028326730243861675
step: 10, loss: 0.004388168454170227
step: 20, loss: 0.0006542751798406243
step: 30, loss: 0.0010410946561023593
step: 40, loss: 0.013204624876379967
step: 50, loss: 0.005083373282104731
step: 60, loss: 0.0012998620513826609
step: 70, loss: 0.0004956320044584572
step: 80, loss: 0.014070714823901653
step: 90, loss: 0.011520590633153915
step: 100, loss: 0.0076753729954361916
step: 110, loss: 0.010834443382918835
step: 120, loss: 0.00017003279936034232
step: 130, loss: 0.028907116502523422
step: 140, loss: 0.010105759836733341
step: 150, loss: 0.008992712944746017
step: 160, loss: 0.0007950814324431121
step: 170, loss: 0.01686490885913372
step: 180, loss: 0.001271918648853898
step: 190, loss: 0.1068849116563797
step: 200, loss: 0.003397110616788268
step: 210, loss: 0.07181580364704132

```

```

step: 220, loss: 0.05125550553202629
step: 230, loss: 0.02818240039050579
step: 240, loss: 0.005438965279608965
step: 250, loss: 0.04109030216932297
step: 260, loss: 0.0014551292406395078
step: 270, loss: 0.0016768694622442126
step: 280, loss: 0.08672686666250229
step: 290, loss: 0.030806973576545715
step: 300, loss: 0.002521953545510769
step: 310, loss: 0.009565997868776321
step: 320, loss: 0.003167359624058008
step: 330, loss: 0.004673016257584095
step: 340, loss: 0.0017325450899079442
step: 350, loss: 0.004488538019359112
step: 360, loss: 0.01444376353174448
step: 370, loss: 0.004207069054245949
step: 380, loss: 0.03357252478599548
step: 390, loss: 0.00022022925259079784
step: 400, loss: 0.009129214100539684
step: 410, loss: 0.0020660676527768373
step: 420, loss: 0.0010909292614087462
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 0.06225867569446564
step: 440, loss: 0.00011543063737917691
step: 450, loss: 0.000566762697417289
step: 460, loss: 0.0001880108320619911
step: 470, loss: 0.1024157926440239
step: 480, loss: 0.15074560046195984
step: 490, loss: 0.004881175700575113
step: 500, loss: 0.0049421763978898525
step: 510, loss: 0.005618422292172909
step: 520, loss: 0.0627681314945221
step: 530, loss: 0.008847671560943127
/usr/local/lib/python3.7/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=0.9431192660550459, f1=0.9441903019213174,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007567986845970154
step: 10, loss: 0.05347031354904175
step: 20, loss: 0.09516283124685287
step: 30, loss: 0.0017189348582178354
step: 40, loss: 0.0008947530295699835
step: 50, loss: 0.0021988076623529196
step: 60, loss: 0.12005147337913513

```

step: 70, loss: 0.00996402557939291
step: 80, loss: 0.0021626665256917477
step: 90, loss: 0.03590674698352814
step: 100, loss: 0.008542451076209545
step: 110, loss: 0.000265309790847823
step: 120, loss: 0.00040912997792474926
step: 130, loss: 0.0022979562636464834
step: 140, loss: 0.0008944349247030914
step: 150, loss: 0.0022813635878264904
step: 160, loss: 0.00022483881912194192
step: 170, loss: 0.014838182367384434
step: 180, loss: 0.0003369531477801502
step: 190, loss: 0.0007374074775725603
step: 200, loss: 0.00040996403549797833
step: 210, loss: 0.021475475281476974
step: 220, loss: 0.0010585528798401356
step: 230, loss: 0.005316411145031452
step: 240, loss: 0.0057873837649822235
step: 250, loss: 0.0036702414508908987
step: 260, loss: 0.00451308349147439
step: 270, loss: 0.00010046570969279855
step: 280, loss: 0.0009028615895658731
step: 290, loss: 0.12922081351280212
step: 300, loss: 0.0018240578938275576
step: 310, loss: 0.0007377158617600799
step: 320, loss: 0.051230743527412415
step: 330, loss: 0.0044743455946445465
step: 340, loss: 0.002100728452205658
step: 350, loss: 0.0006720445235259831
step: 360, loss: 0.01531631126999855
step: 370, loss: 0.0014543691650032997
step: 380, loss: 0.0001187785601359792
step: 390, loss: 0.0001985111739486456
step: 400, loss: 0.007113984320312738
step: 410, loss: 0.017641467973589897
step: 420, loss: 0.004743541125208139
step: 430, loss: 0.004931422881782055
step: 440, loss: 0.0006519725429825485
step: 450, loss: 0.0012735810596495867
step: 460, loss: 0.0021490692161023617
step: 470, loss: 0.020099235698580742
step: 480, loss: 0.08491106331348419
step: 490, loss: 0.001250449800863862
step: 500, loss: 0.0041560279205441475
step: 510, loss: 0.01867905631661415
step: 520, loss: 0.0007766042836010456
step: 530, loss: 0.04180985316634178
/usr/local/lib/python3.7/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=0.9423604757548033, f1=0.9427272727272727,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004192752297967672
step: 10, loss: 0.005346923600882292
step: 20, loss: 0.06501850485801697
step: 30, loss: 0.00026799971237778664
step: 40, loss: 0.0014098025858402252
step: 50, loss: 0.06988012790679932
step: 60, loss: 0.000214589003007859
step: 70, loss: 0.0001761794410413131
step: 80, loss: 0.0005561114521697164
step: 90, loss: 0.0003462099703028798
step: 100, loss: 0.0002909491886384785
step: 110, loss: 0.0007969379075802863
step: 120, loss: 0.00022622194956056774
step: 130, loss: 0.03129217401146889
step: 140, loss: 0.0013171132886782289
step: 150, loss: 0.0014891093596816063
step: 160, loss: 0.0020917244255542755
step: 170, loss: 0.0003283658588770777
step: 180, loss: 0.00017026888963300735
step: 190, loss: 0.00032996671507135034
step: 200, loss: 0.0003811245842371136
step: 210, loss: 0.00019734221859835088
step: 220, loss: 0.00035672736703418195
step: 230, loss: 0.00015000942221377045
step: 240, loss: 0.08702774345874786
step: 250, loss: 0.0006154726725071669
step: 260, loss: 6.14183591096662e-05
step: 270, loss: 0.05408139154314995
step: 280, loss: 0.03178141266107559
step: 290, loss: 0.00018215834279544652
step: 300, loss: 0.0007885849918238819
step: 310, loss: 0.00032092456240206957
step: 320, loss: 0.00034799304557964206
step: 330, loss: 0.0001977062929654494
step: 340, loss: 0.10246788710355759
step: 350, loss: 0.0008485498838126659
step: 360, loss: 0.013723563402891159
step: 370, loss: 0.0013210217002779245
step: 380, loss: 0.00012720143422484398
step: 390, loss: 0.001720127183943987
step: 400, loss: 0.0006148135289549828
```

```

step: 410, loss: 0.0010610512690618634
step: 420, loss: 8.877993241185322e-05
step: 430, loss: 0.00072203524177894
step: 440, loss: 0.0008570700883865356
step: 450, loss: 0.053234051913022995
step: 460, loss: 6.427209882531315e-05
step: 470, loss: 0.0002076432720059529
step: 480, loss: 0.0022077751345932484
step: 490, loss: 0.009049305692315102
step: 500, loss: 0.00030009428155608475
step: 510, loss: 0.0037166294641792774
step: 520, loss: 0.0001881768403109163
step: 530, loss: 0.007006017956882715
/usr/local/lib/python3.7/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=0.9348946135831382, f1=0.9368029739776952,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038439666968770325
step: 10, loss: 0.00034540804335847497
step: 20, loss: 0.0009278280194848776
step: 30, loss: 0.0005408489378169179
step: 40, loss: 0.0009123117779381573
step: 50, loss: 0.0025996456388384104
step: 60, loss: 0.003979658707976341
step: 70, loss: 0.008833732455968857
step: 80, loss: 0.014789308421313763
step: 90, loss: 0.035614416003227234
step: 100, loss: 0.18184232711791992
step: 110, loss: 0.0004690089845098555
step: 120, loss: 0.00028767448384314775
step: 130, loss: 0.016452457755804062
step: 140, loss: 0.0007400819449685514
step: 150, loss: 0.006613555829972029
step: 160, loss: 5.530329144676216e-05
step: 170, loss: 0.00024865695741027594
step: 180, loss: 0.0010233456268906593
step: 190, loss: 0.0034546824172139168
step: 200, loss: 0.0011970741907134652
step: 210, loss: 9.299257362727076e-05
step: 220, loss: 0.0001241346908500418
step: 230, loss: 0.003955059684813023
step: 240, loss: 0.0008032615296542645
step: 250, loss: 0.0006898937281221151
step: 260, loss: 0.00011342236393829808

```

```

step: 270, loss: 0.14872992038726807
step: 280, loss: 0.0002787074772641063
step: 290, loss: 0.0005318389157764614
step: 300, loss: 0.002928441157564521
step: 310, loss: 0.000834379403386265
step: 320, loss: 0.008503139019012451
step: 330, loss: 0.0015341945691034198
step: 340, loss: 0.03697490319609642
step: 350, loss: 0.054751139134168625
step: 360, loss: 0.00566209526732564
step: 370, loss: 0.000844664522446692
step: 380, loss: 0.0008791991858743131
step: 390, loss: 0.00025744730373844504
step: 400, loss: 0.0018557131988927722
step: 410, loss: 0.0014105531154200435
step: 420, loss: 0.0024853511713445187
step: 430, loss: 0.0017913669580593705
step: 440, loss: 0.00027497121482156217
step: 450, loss: 0.00038169484469108284
step: 460, loss: 0.000790067482739687
step: 470, loss: 0.12008503079414368
step: 480, loss: 0.04279835894703865
step: 490, loss: 0.00025778712006285787
step: 500, loss: 0.0015667215920984745
step: 510, loss: 0.017481178045272827
step: 520, loss: 0.007936621084809303
step: 530, loss: 0.012760164216160774
/usr/local/lib/python3.7/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=0.9435559736594544, f1=0.9449112978524743,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003525524807628244
step: 10, loss: 0.028123067691922188
step: 20, loss: 0.003930854611098766
step: 30, loss: 0.0005732719437219203
step: 40, loss: 7.775230187689885e-05
step: 50, loss: 0.001615738496184349
step: 60, loss: 0.024580052122473717
step: 70, loss: 9.848659829003736e-05
step: 80, loss: 0.0008668515947647393
step: 90, loss: 0.001130554242990911
step: 100, loss: 0.00010292127262800932
step: 110, loss: 0.0002032196061918512
step: 120, loss: 0.0018105354392901063

```

```

step: 130, loss: 0.010472914204001427
step: 140, loss: 0.0008421468664892018
step: 150, loss: 0.0005240549216978252
step: 160, loss: 0.030515236780047417
step: 170, loss: 0.01220968272536993
step: 180, loss: 9.302154649049044e-05
step: 190, loss: 0.0013986020348966122
step: 200, loss: 0.0015691723674535751
step: 210, loss: 0.022515881806612015
step: 220, loss: 0.00048075782251544297
step: 230, loss: 4.236818494973704e-05
step: 240, loss: 0.0002194448170484975
step: 250, loss: 6.055185440345667e-05
step: 260, loss: 3.841351644950919e-05
step: 270, loss: 0.001571480417624116
step: 280, loss: 0.04226793348789215
step: 290, loss: 9.058790601557121e-05
step: 300, loss: 0.01005515269935131
step: 310, loss: 0.00021561171161010861
step: 320, loss: 0.012453639879822731
step: 330, loss: 0.00030300579965114594
step: 340, loss: 8.447960135526955e-05
step: 350, loss: 0.00011901209654752165
step: 360, loss: 0.10433193296194077
step: 370, loss: 0.00013564729306381196
step: 380, loss: 0.0015297782374545932
step: 390, loss: 0.11120890825986862
step: 400, loss: 7.747753261355683e-05
step: 410, loss: 0.0005688324454240501
step: 420, loss: 0.004020680673420429
step: 430, loss: 0.04402010142803192
step: 440, loss: 0.0008488519233651459
step: 450, loss: 0.00011483138951007277
step: 460, loss: 0.0004920171340927482
step: 470, loss: 0.0009264570544473827
step: 480, loss: 0.009427216835319996
step: 490, loss: 0.00036680224002338946
step: 500, loss: 0.00012831062485929579
step: 510, loss: 8.447307482128963e-05
step: 520, loss: 0.0019802909810096025
step: 530, loss: 0.01305887196213007
/usr/local/lib/python3.7/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=0.943502824858757, f1=0.9445221445221446,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020426749251782894
step: 10, loss: 0.0017951823538169265
step: 20, loss: 0.003966579679399729
step: 30, loss: 0.00010172255133511499
step: 40, loss: 6.549743557116017e-05
step: 50, loss: 2.7584856070461683e-05
step: 60, loss: 3.834970993921161e-05
step: 70, loss: 7.95487649156712e-05
step: 80, loss: 0.001367076183669269
step: 90, loss: 3.462542736087926e-05
step: 100, loss: 8.577038533985615e-05
step: 110, loss: 3.523576015140861e-05
step: 120, loss: 5.5960659665288404e-05
step: 130, loss: 0.0009607035899534822
step: 140, loss: 7.220516272354871e-05
step: 150, loss: 7.972119055921212e-05
step: 160, loss: 0.0021165241487324238
step: 170, loss: 0.006125062704086304
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 180, loss: 3.354405998834409e-05
step: 190, loss: 3.305647260276601e-05
step: 200, loss: 0.00010083545203087851
step: 210, loss: 0.00019200843235012144
step: 220, loss: 0.0004651309282053262
step: 230, loss: 2.784932621580083e-05
step: 240, loss: 4.666428139898926e-05
step: 250, loss: 9.126473742071539e-05
step: 260, loss: 0.008545349352061749
step: 270, loss: 2.0015651898575015e-05
step: 280, loss: 4.780710150953382e-05
step: 290, loss: 0.05027167499065399
step: 300, loss: 7.118321809684858e-05
step: 310, loss: 0.02870832569897175
step: 320, loss: 0.0009251062874682248
step: 330, loss: 0.000791500206105411
step: 340, loss: 0.07227783650159836
step: 350, loss: 0.0037646866403520107
step: 360, loss: 0.0005079603870399296
step: 370, loss: 0.0036344241816550493
step: 380, loss: 0.00034387840423732996
step: 390, loss: 0.00030576041899621487
step: 400, loss: 0.006378231570124626
step: 410, loss: 7.176766666816548e-05
step: 420, loss: 0.0005546488682739437
step: 430, loss: 7.098046626197174e-05
step: 440, loss: 0.0009128281380981207
step: 450, loss: 6.798231333959848e-05
```



```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 460, loss: 0.00013214688806328923
step: 470, loss: 4.871035707765259e-05
step: 480, loss: 9.87806633929722e-05
step: 490, loss: 0.0003806966997217387
step: 500, loss: 0.0012742956168949604
step: 510, loss: 0.00016809575026854873
step: 520, loss: 0.00015239643107634038
step: 530, loss: 5.704408249584958e-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=0.9389277389277391, f1=0.9448307834955958,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001331572129856795
step: 10, loss: 4.018356048618443e-05
step: 20, loss: 3.848879714496434e-05
step: 30, loss: 5.227707515587099e-05
step: 40, loss: 6.16046236245893e-05
step: 50, loss: 0.00019895777222700417
step: 60, loss: 0.00010442228813190013
step: 70, loss: 8.819897630019113e-05
step: 80, loss: 2.7257219699095003e-05
step: 90, loss: 2.9876086045987904e-05
step: 100, loss: 3.144798392895609e-05
step: 110, loss: 2.188201506214682e-05
step: 120, loss: 0.018155761063098907
step: 130, loss: 0.0001432271092198789
step: 140, loss: 0.007534248288720846
step: 150, loss: 4.8377107304986566e-05
step: 160, loss: 2.7457996111479588e-05
step: 170, loss: 0.00017437951464671642
step: 180, loss: 6.991508416831493e-05
step: 190, loss: 0.0002739575575105846
step: 200, loss: 0.0013731238432228565
step: 210, loss: 0.00013159203808754683
step: 220, loss: 0.006650768220424652
step: 230, loss: 0.0015422261785715818
step: 240, loss: 0.00010824314813362435
step: 250, loss: 4.244687806931324e-05
step: 260, loss: 0.0008784852107055485
step: 270, loss: 0.0009230792638845742
step: 280, loss: 9.173254511551932e-05
step: 290, loss: 2.7167969165020622e-05
step: 300, loss: 3.796193050220609e-05

```

```

step: 310, loss: 0.00033354287734255195
step: 320, loss: 0.002574209589511156
step: 330, loss: 0.0006979400059208274
step: 340, loss: 0.013597637414932251
step: 350, loss: 0.00011017148790415376
step: 360, loss: 4.346649438957684e-05
step: 370, loss: 0.007930207997560501
step: 380, loss: 0.00013902240607421845
step: 390, loss: 9.565043001202866e-05
step: 400, loss: 0.0010450505651533604
step: 410, loss: 0.0026807901449501514
step: 420, loss: 3.8077454519225284e-05
step: 430, loss: 2.995285831275396e-05
step: 440, loss: 0.00024575230781920254
step: 450, loss: 2.5774734240258113e-05
step: 460, loss: 7.934540190035477e-05
step: 470, loss: 7.100482616806403e-05
step: 480, loss: 3.3190222893608734e-05
step: 490, loss: 0.00021598723833449185
step: 500, loss: 0.0010118552017956972
step: 510, loss: 2.447841143293772e-05
step: 520, loss: 6.469246000051498e-05
step: 530, loss: 0.0015547149814665318
/usr/local/lib/python3.7/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=0.9436422915696321, f1=0.9480037140204272,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011054299829993397
step: 10, loss: 7.309818465728313e-05
step: 20, loss: 0.0003622969670686871
step: 30, loss: 0.00020595741807483137
step: 40, loss: 0.0013810155214741826
step: 50, loss: 0.00025956748868338764
step: 60, loss: 0.027228888124227524
step: 70, loss: 6.714541814289987e-05
step: 80, loss: 4.7438148612855e-05
step: 90, loss: 2.8032045520376414e-05
step: 100, loss: 0.002362996805459261
step: 110, loss: 9.315713396063074e-05
step: 120, loss: 0.0001291505031986162
step: 130, loss: 2.577466148068197e-05
step: 140, loss: 3.515679418342188e-05
step: 150, loss: 1.9184873963240534e-05
step: 160, loss: 0.00011562931467778981

```

```

step: 170, loss: 4.0070375689538196e-05
step: 180, loss: 0.0002821396046783775
step: 190, loss: 0.0018559733871370554
step: 200, loss: 0.0016853543929755688
step: 210, loss: 0.00019557894847821444
step: 220, loss: 0.0029475532937794924
step: 230, loss: 5.1553884986788034e-05
step: 240, loss: 0.0005900523974560201
step: 250, loss: 3.17944104608614e-05
step: 260, loss: 4.070651266374625e-05
step: 270, loss: 0.0019011043477803469
step: 280, loss: 0.0003893358225468546
step: 290, loss: 0.0008223338518291712
step: 300, loss: 0.0003073445986956358
step: 310, loss: 7.035864109639078e-05
step: 320, loss: 0.0007529630674980581
step: 330, loss: 0.0001099786750273779
step: 340, loss: 0.0004912807489745319
step: 350, loss: 0.0001786801149137318
step: 360, loss: 2.1990026652929373e-05
step: 370, loss: 0.0016973083838820457
step: 380, loss: 4.440459088073112e-05
step: 390, loss: 3.130139521090314e-05
step: 400, loss: 3.81035752070602e-05
step: 410, loss: 0.0005750021664425731
step: 420, loss: 0.00010894893057411537
step: 430, loss: 3.6368524888530374e-05
step: 440, loss: 0.00048486993182450533
step: 450, loss: 0.0013315679971128702
step: 460, loss: 0.00017196261615026742
step: 470, loss: 0.00021352637850213796
step: 480, loss: 0.001483640051446855
step: 490, loss: 0.0004970543668605387
step: 500, loss: 0.00016953122394625098
step: 510, loss: 9.99880867311731e-05
step: 520, loss: 0.0005334315937943757
step: 530, loss: 0.0004339656443335116
/usr/local/lib/python3.7/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=0.9422263973696571, f1=0.9415067852129153,
best_f1=0.939435968562182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.102343154954724e-05
step: 10, loss: 3.726759314304218e-05
step: 20, loss: 0.0011863451218232512

```

step: 30, loss: 0.003216284094378352
step: 40, loss: 0.0001225369342137128
step: 50, loss: 0.054624635726213455
step: 60, loss: 0.0013250766787678003
step: 70, loss: 5.2537230658344924e-05
step: 80, loss: 3.258653305238113e-05
step: 90, loss: 0.0001543032267363742
step: 100, loss: 0.0012567657977342606
step: 110, loss: 3.842053047264926e-05
step: 120, loss: 3.6727633414557204e-05
step: 130, loss: 0.0017510028555989265
step: 140, loss: 6.696212949464098e-05
step: 150, loss: 6.728097650920972e-05
step: 160, loss: 2.0224242689437233e-05
step: 170, loss: 5.2162711654091254e-05
step: 180, loss: 7.35795620130375e-05
step: 190, loss: 3.32951130985748e-05
step: 200, loss: 9.326326107839122e-05
step: 210, loss: 0.0003480332379695028
step: 220, loss: 6.0828715504612774e-05
step: 230, loss: 1.94120602827752e-05
step: 240, loss: 8.414554031332955e-05
step: 250, loss: 1.9129050997435115e-05
step: 260, loss: 3.5132583434460685e-05
step: 270, loss: 0.0004936957266181707
step: 280, loss: 0.0023095342330634594
step: 290, loss: 0.00026184471789747477
step: 300, loss: 0.00046306452713906765
step: 310, loss: 3.871730950777419e-05
step: 320, loss: 2.26633910642704e-05
step: 330, loss: 4.021905624540523e-05
step: 340, loss: 2.748816041275859e-05
step: 350, loss: 0.00035127237788401544
step: 360, loss: 0.0005947224562987685
step: 370, loss: 0.001615645713172853
step: 380, loss: 0.00016813284310046583
step: 390, loss: 0.00015336195065174252
step: 400, loss: 2.683221100596711e-05
step: 410, loss: 0.0018608203390613198
step: 420, loss: 0.006556607782840729
step: 430, loss: 0.008836396969854832
step: 440, loss: 0.0004921305226162076
step: 450, loss: 4.022604844067246e-05
step: 460, loss: 1.7057887816918083e-05
step: 470, loss: 1.8562659533927217e-05
step: 480, loss: 0.0004740249423775822
step: 490, loss: 0.00016455953300464898
step: 500, loss: 6.246419798117131e-05

```

step: 510, loss: 0.0008099854458123446
step: 520, loss: 1.9758608686970547e-05
step: 530, loss: 2.835953637259081e-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 12: dev_f1=0.9464867380176826, f1=0.9450651769087525,
best_f1=0.9450651769087525
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020250135276000947
step: 10, loss: 5.9489673731150106e-05
step: 20, loss: 1.7061556718545035e-05
step: 30, loss: 3.427075353101827e-05
step: 40, loss: 1.7493654013378546e-05
step: 50, loss: 6.60140867694281e-05
step: 60, loss: 2.181058698624838e-05
step: 70, loss: 0.00018507830100134015
step: 80, loss: 3.979688335675746e-05
step: 90, loss: 0.02490376867353916
step: 100, loss: 5.6120748922694474e-05
step: 110, loss: 0.00011585757602006197
step: 120, loss: 5.2049079386051744e-05
step: 130, loss: 2.5058212486328557e-05
step: 140, loss: 2.89662966679316e-05
step: 150, loss: 7.701788854319602e-05
step: 160, loss: 3.1490184483118355e-05
step: 170, loss: 0.000295746314805001
step: 180, loss: 0.0004537857894320041
step: 190, loss: 2.2172102035256103e-05
step: 200, loss: 2.3747988961986266e-05
step: 210, loss: 3.885419209836982e-05
step: 220, loss: 1.668528602749575e-05
step: 230, loss: 1.5705625628470443e-05
step: 240, loss: 0.0001708838390186429
step: 250, loss: 1.3723846677748952e-05
step: 260, loss: 1.831317786127329e-05
step: 270, loss: 1.648031138756778e-05
step: 280, loss: 2.2905920559423976e-05
step: 290, loss: 0.0001606843143235892
step: 300, loss: 1.9047130990657024e-05
step: 310, loss: 0.00016050264821387827
step: 320, loss: 0.00483768479898572

```

```

step: 330, loss: 0.001256215269677341
step: 340, loss: 2.30098648899002e-05
step: 350, loss: 2.164713805541396e-05
step: 360, loss: 1.739687468216289e-05
step: 370, loss: 0.0028925302904099226
step: 380, loss: 2.7692014555213973e-05
step: 390, loss: 0.0001571920729475096
step: 400, loss: 0.00027438325923867524
step: 410, loss: 0.000904412823729217
step: 420, loss: 2.097274591505993e-05
step: 430, loss: 0.00022211187751963735
step: 440, loss: 3.17763006023597e-05
step: 450, loss: 3.750581163330935e-05
step: 460, loss: 0.0012670353753492236
step: 470, loss: 1.3492806829162873e-05
step: 480, loss: 4.003447975264862e-05
step: 490, loss: 2.154634967155289e-05
step: 500, loss: 3.7923178751952946e-05
step: 510, loss: 6.47998895146884e-05
step: 520, loss: 0.00029454808100126684
step: 530, loss: 1.91288381756749e-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=0.9456928838951311, f1=0.9430970149253732,
best_f1=0.9450651769087525
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011885642306879163
step: 10, loss: 1.7772277715266682e-05
step: 20, loss: 6.457250128732994e-05
step: 30, loss: 1.3660523109138012e-05
step: 40, loss: 3.6381257814355195e-05
step: 50, loss: 7.089573773555458e-05
step: 60, loss: 1.3068155567452777e-05
step: 70, loss: 1.6353456885553896e-05
step: 80, loss: 5.921438423683867e-05
step: 90, loss: 1.4927067240932956e-05
step: 100, loss: 4.415628427523188e-05
step: 110, loss: 1.802994120225776e-05
step: 120, loss: 1.1898482625838369e-05
step: 130, loss: 0.023889897391200066
step: 140, loss: 0.0002951112110167742
step: 150, loss: 1.0505246791581158e-05
step: 160, loss: 7.090721192071214e-05
step: 170, loss: 4.819529567612335e-05
step: 180, loss: 1.2274731489014812e-05

```

```

step: 190, loss: 9.007690823636949e-05
step: 200, loss: 0.0010623778216540813
step: 210, loss: 1.407399759045802e-05
step: 220, loss: 1.3883976862416603e-05
step: 230, loss: 3.246480628149584e-05
step: 240, loss: 1.3403460798144806e-05
step: 250, loss: 3.5426681279204786e-05
step: 260, loss: 0.00028776060207746923
step: 270, loss: 1.4982945685915183e-05
step: 280, loss: 1.4949382602935657e-05
step: 290, loss: 0.0004186325240880251
step: 300, loss: 2.385962761763949e-05
step: 310, loss: 2.6179332053288817e-05
step: 320, loss: 0.0001822734484449029
step: 330, loss: 4.317864659242332e-05
step: 340, loss: 1.920725844684057e-05
step: 350, loss: 0.000150249368743971
step: 360, loss: 6.576617306564003e-05
step: 370, loss: 2.2510404960485175e-05
step: 380, loss: 2.1318739527487196e-05
step: 390, loss: 0.02312825433909893
step: 400, loss: 9.77837698883377e-05
step: 410, loss: 1.838733624026645e-05
step: 420, loss: 1.661429814703297e-05
step: 430, loss: 1.822000922402367e-05
step: 440, loss: 6.815144297434017e-05
step: 450, loss: 2.4239889171440154e-05
step: 460, loss: 0.0003588833787944168
step: 470, loss: 1.3641835721500684e-05
step: 480, loss: 1.4737070159753785e-05
step: 490, loss: 1.913979576784186e-05
step: 500, loss: 1.2431186405592598e-05
step: 510, loss: 2.5923287466866896e-05
step: 520, loss: 0.0005861318204551935
step: 530, loss: 1.4710973118781112e-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=0.9449626865671642, f1=0.9418334108887855,
best_f1=0.9450651769087525
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3501163013861515e-05
step: 10, loss: 1.5802421330590732e-05
step: 20, loss: 2.048414717137348e-05
step: 30, loss: 6.335578655125573e-05
step: 40, loss: 0.08915434032678604

```

step: 50, loss: 0.00029375290614552796
step: 60, loss: 1.1108730177511461e-05
step: 70, loss: 4.588376032188535e-05
step: 80, loss: 4.613583951140754e-05
step: 90, loss: 1.35897244035732e-05
step: 100, loss: 2.731943823164329e-05
step: 110, loss: 0.001800857251510024
step: 120, loss: 7.374511915259063e-05
step: 130, loss: 0.09912941604852676
step: 140, loss: 2.0309467799961567e-05
step: 150, loss: 1.5090989109012298e-05
step: 160, loss: 7.406144868582487e-05
step: 170, loss: 1.1946901395276655e-05
step: 180, loss: 1.9544791939551942e-05
step: 190, loss: 3.328813909320161e-05
step: 200, loss: 1.3593437870440539e-05
step: 210, loss: 0.00012895188410766423
step: 220, loss: 1.0654259313014336e-05
step: 230, loss: 1.7415208276361227e-05
step: 240, loss: 1.9262772184447385e-05
step: 250, loss: 2.0287150618969463e-05
step: 260, loss: 2.98467402899405e-05
step: 270, loss: 1.8551178072812036e-05
step: 280, loss: 1.3351287634577602e-05
step: 290, loss: 1.1563209227460902e-05
step: 300, loss: 1.323582819168223e-05
step: 310, loss: 4.155136412009597e-05
step: 320, loss: 0.00021162502525839955
step: 330, loss: 2.6357702154200524e-05
step: 340, loss: 2.5359788196510635e-05
step: 350, loss: 1.0598377230053302e-05
step: 360, loss: 0.002089268760755658
step: 370, loss: 1.491187413193984e-05
step: 380, loss: 1.9519829947967082e-05
step: 390, loss: 1.794071613403503e-05
step: 400, loss: 0.00015552843979094177
step: 410, loss: 1.608927414054051e-05
step: 420, loss: 2.553143713157624e-05
step: 430, loss: 0.0012356694787740707
step: 440, loss: 7.7886572398711e-05
step: 450, loss: 1.1719657777575776e-05
step: 460, loss: 1.4479929632216226e-05
step: 470, loss: 0.0042153759859502316
step: 480, loss: 1.0255653251078911e-05
step: 490, loss: 1.0803267286974005e-05
step: 500, loss: 4.50301195087377e-05
step: 510, loss: 1.3638153177453205e-05
step: 520, loss: 1.1600470315897837e-05


```

step: 530, loss: 0.00022545513638760895
/usr/local/lib/python3.7/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=0.9446254071661238, f1=0.9434137291280148,
best_f1=0.9450651769087525
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, 191.77it/s]
load_f1 = 0.9464867380176826
real_f1 = 0.9460465116279071
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, 185.10it/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: 0.554735004901886
step: 10, loss: 0.38532599806785583
step: 20, loss: 0.38432571291923523
step: 30, loss: 0.33126479387283325
step: 40, loss: 0.19463521242141724
step: 50, loss: 0.3984035849571228
step: 60, loss: 0.24744853377342224
step: 70, loss: 0.15167337656021118
step: 80, loss: 0.20663009583950043
step: 90, loss: 0.2966393232345581
step: 100, loss: 0.33188292384147644
step: 110, loss: 0.17780068516731262
step: 120, loss: 0.17988643050193787
step: 130, loss: 0.17180994153022766
step: 140, loss: 0.20685666799545288
step: 150, loss: 0.2029298096895218
step: 160, loss: 0.32477006316185
step: 170, loss: 0.23641721904277802
step: 180, loss: 0.11738386005163193
step: 190, loss: 0.25506874918937683
step: 200, loss: 0.23002296686172485
step: 210, loss: 0.24380351603031158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6278195488721804, f1=0.7029126213592233,
best_f1=0.7029126213592233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07689442485570908
step: 10, loss: 0.1827174574136734
step: 20, loss: 0.19886235892772675
step: 30, loss: 0.22453539073467255
step: 40, loss: 0.12766216695308685
step: 50, loss: 0.133833646774292
step: 60, loss: 0.31418949365615845
step: 70, loss: 0.15164370834827423
step: 80, loss: 0.17680898308753967
step: 90, loss: 0.13935479521751404
step: 100, loss: 0.008898114785552025
step: 110, loss: 0.07502800226211548
step: 120, loss: 0.134477898478508
step: 130, loss: 0.013658873736858368
step: 140, loss: 0.16101397573947906
step: 150, loss: 0.16577006876468658
step: 160, loss: 0.2399263232946396
step: 170, loss: 0.10658011585474014
step: 180, loss: 0.1935301274061203
step: 190, loss: 0.2915789783000946
step: 200, loss: 0.057477742433547974
step: 210, loss: 0.15111391246318817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6455696202531644, f1=0.7027027027027027,
best_f1=0.7027027027027027
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.062026623636484146
step: 10, loss: 0.13229502737522125
step: 20, loss: 0.11183695495128632
step: 30, loss: 0.10834002494812012
step: 40, loss: 0.07118631154298782
step: 50, loss: 0.1141585186123848
step: 60, loss: 0.059693191200494766
step: 70, loss: 0.02720778062939644
step: 80, loss: 0.14189498126506805

```

```

step: 90, loss: 0.055083561688661575
step: 100, loss: 0.09779741615056992
step: 110, loss: 0.3984386920928955
step: 120, loss: 0.15508483350276947
step: 130, loss: 0.11063193529844284
step: 140, loss: 0.09306497871875763
step: 150, loss: 0.20221492648124695
step: 160, loss: 0.01023570355027914
step: 170, loss: 0.037610530853271484
step: 180, loss: 0.13024768233299255
step: 190, loss: 0.21329574286937714
step: 200, loss: 0.05647855997085571
step: 210, loss: 0.11167080700397491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.691358024691358,
best_f1=0.691358024691358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07685039192438126
step: 10, loss: 0.050522416830062866
step: 20, loss: 0.037265975028276443
step: 30, loss: 0.10249260067939758
step: 40, loss: 0.011888034641742706
step: 50, loss: 0.09347593784332275
step: 60, loss: 0.08738745748996735
step: 70, loss: 0.13098955154418945
step: 80, loss: 0.05192563310265541
step: 90, loss: 0.013022358529269695
step: 100, loss: 0.173015296459198
step: 110, loss: 0.1454635113477707
step: 120, loss: 0.046405404806137085
step: 130, loss: 0.14756987988948822
step: 140, loss: 0.11074388772249222
step: 150, loss: 0.010100050829350948
step: 160, loss: 0.029254259541630745
step: 170, loss: 0.12027069926261902
step: 180, loss: 0.2619548738002777
step: 190, loss: 0.04275379702448845
step: 200, loss: 0.09708643704652786
step: 210, loss: 0.3225546181201935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6835937499999999, 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: 0.09019454568624496
step: 10, loss: 0.04592903330922127
step: 20, loss: 0.071884885430336
step: 30, loss: 0.012610363774001598
step: 40, loss: 0.04469282180070877
step: 50, loss: 0.030911238864064217
step: 60, loss: 0.0350593663752079
step: 70, loss: 0.006692846305668354
step: 80, loss: 0.04410182684659958
step: 90, loss: 0.11468271166086197
step: 100, loss: 0.007263215258717537
step: 110, loss: 0.11767423897981644
step: 120, loss: 0.06645341962575912
step: 130, loss: 0.062266863882541656
step: 140, loss: 0.02646504156291485
step: 150, loss: 0.061213795095682144
step: 160, loss: 0.10680900514125824
step: 170, loss: 0.05141142010688782
step: 180, loss: 0.04864848777651787
step: 190, loss: 0.034291062504053116
step: 200, loss: 0.059707917273044586
step: 210, loss: 0.012910725548863411
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7024952015355086, f1=0.6982922201138521,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04459589719772339
step: 10, loss: 0.035277243703603745
step: 20, loss: 0.06707737594842911

```

```

step: 30, loss: 0.001572235720232129
step: 40, loss: 0.001431611948646605
step: 50, loss: 0.022777052596211433
step: 60, loss: 0.055781129747629166
step: 70, loss: 0.015372870489954948
step: 80, loss: 0.10551510751247406
step: 90, loss: 0.12661787867546082
step: 100, loss: 0.01540947612375021
step: 110, loss: 0.005099620204418898
step: 120, loss: 0.12512382864952087
step: 130, loss: 0.05659735947847366
step: 140, loss: 0.1093391627073288
step: 150, loss: 0.013802996836602688
step: 160, loss: 0.012467347085475922
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.06738471239805222
step: 180, loss: 0.0195819903165102
step: 190, loss: 0.15725606679916382
step: 200, loss: 0.004472676664590836
step: 210, loss: 0.020778337493538857
/usr/local/lib/python3.7/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=0.6741154562383612, f1=0.6861313868613138,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004565644543617964
step: 10, loss: 0.007798554375767708
step: 20, loss: 0.029466450214385986
step: 30, loss: 0.007293494418263435
step: 40, loss: 0.06307201832532883
step: 50, loss: 0.07840446382761002
step: 60, loss: 0.014733031392097473
step: 70, loss: 0.03272585570812225
step: 80, loss: 0.03128219395875931
step: 90, loss: 0.0027973493561148643
step: 100, loss: 0.004149234853684902
step: 110, loss: 0.12488626688718796
step: 120, loss: 0.03187492489814758
step: 130, loss: 0.01422555185854435
step: 140, loss: 0.014417780563235283
step: 150, loss: 0.0220321211963892
step: 160, loss: 0.026536835357546806
step: 170, loss: 0.003910951782017946
step: 180, loss: 0.11490121483802795
step: 190, loss: 0.10436524450778961

```

```

step: 200, loss: 0.001847187289968133
step: 210, loss: 0.0074443272314965725
/usr/local/lib/python3.7/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=0.6836935166994106, f1=0.6946564885496184,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00643228180706501
step: 10, loss: 0.04676388204097748
step: 20, loss: 0.0017199417343363166
step: 30, loss: 0.13951969146728516
step: 40, loss: 0.0025759260170161724
step: 50, loss: 0.0019609692972153425
step: 60, loss: 0.10809601843357086
step: 70, loss: 0.053991008549928665
step: 80, loss: 0.02082185447216034
step: 90, loss: 0.033788155764341354
step: 100, loss: 0.025624359026551247
step: 110, loss: 0.03976665064692497
step: 120, loss: 0.000872071657795459
step: 130, loss: 0.0015073335962370038
step: 140, loss: 0.04641430452466011
step: 150, loss: 0.003901669755578041
step: 160, loss: 0.02558823488652706
step: 170, loss: 0.007638496812433004
step: 180, loss: 0.040360890328884125
step: 190, loss: 0.010249379090964794
step: 200, loss: 0.0022730363998562098
step: 210, loss: 0.20285958051681519
/usr/local/lib/python3.7/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=0.6936170212765959, f1=0.7058823529411764,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02054341323673725
step: 10, loss: 0.008128275163471699
step: 20, loss: 0.007229821756482124
step: 30, loss: 0.00986101757735014
step: 40, loss: 0.0347636416554451
step: 50, loss: 0.07975010573863983
step: 60, loss: 0.11358565837144852
step: 70, loss: 0.03158752620220184

```



```

step: 80, loss: 0.001917649875395
step: 90, loss: 0.00027359116938896477
step: 100, loss: 0.009424306452274323
step: 110, loss: 0.024637624621391296
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.04162197932600975
step: 130, loss: 0.003121413057669997
step: 140, loss: 0.0045874156057834625
step: 150, loss: 0.04559152573347092
step: 160, loss: 0.0002918066456913948
step: 170, loss: 0.011483718641102314
step: 180, loss: 0.07349257916212082
step: 190, loss: 0.007637652102857828
step: 200, loss: 0.04170048609375954
step: 210, loss: 0.002305375412106514
/usr/local/lib/python3.7/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=0.6652806652806652, f1=0.7014028056112224,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01107748132199049
step: 10, loss: 0.030360177159309387
step: 20, loss: 0.0018430408090353012
step: 30, loss: 0.043173130601644516
step: 40, loss: 0.10215617716312408
step: 50, loss: 0.12203021347522736
step: 60, loss: 0.005481828469783068
step: 70, loss: 0.12381631880998611
step: 80, loss: 0.015543009154498577
step: 90, loss: 0.013910632580518723
step: 100, loss: 0.001612001215107739
step: 110, loss: 0.0037009783554822206
step: 120, loss: 0.002498059533536434
step: 130, loss: 0.032745420932769775
step: 140, loss: 0.004099245183169842
step: 150, loss: 0.03354334086179733
step: 160, loss: 0.003071555867791176
step: 170, loss: 0.0016043893992900848
step: 180, loss: 0.006312901619821787
step: 190, loss: 0.09118853509426117
step: 200, loss: 0.0006222198135219514
step: 210, loss: 0.031073281541466713
/usr/local/lib/python3.7/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=0.6695464362850972, f1=0.7027027027027027,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016164233908057213
step: 10, loss: 0.02514050528407097
step: 20, loss: 0.021411068737506866
step: 30, loss: 0.0009517583530396223
step: 40, loss: 0.004281703848391771
step: 50, loss: 0.019709588959813118
step: 60, loss: 0.003728017210960388
step: 70, loss: 0.008301686495542526
step: 80, loss: 0.008483569137752056
step: 90, loss: 0.00018036190886050463
step: 100, loss: 0.01162766758352518
step: 110, loss: 0.012606017291545868
step: 120, loss: 0.024502675980329514
step: 130, loss: 0.018854819238185883
step: 140, loss: 7.598228694405407e-05
step: 150, loss: 0.005932766478508711
step: 160, loss: 0.007924076169729233
step: 170, loss: 0.06388776749372482
step: 180, loss: 0.000935536518227309
step: 190, loss: 0.006497329566627741
step: 200, loss: 0.0004918287741020322
step: 210, loss: 0.000291334989015013
/usr/local/lib/python3.7/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=0.6768558951965066, f1=0.721868365180467,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09785252064466476
step: 10, loss: 0.0003797124663833529
step: 20, loss: 0.005002535413950682
step: 30, loss: 0.0010478099575266242
step: 40, loss: 0.0029622025322169065
step: 50, loss: 0.0006818770198151469
step: 60, loss: 0.0011768895201385021
step: 70, loss: 0.02428152970969677
step: 80, loss: 0.000157211979967542
step: 90, loss: 0.03435467556118965
step: 100, loss: 0.01567014865577221
step: 110, loss: 0.00396962882950902
step: 120, loss: 0.00048315717140212655

```

```

step: 130, loss: 0.0013717759866267443
step: 140, loss: 0.00012080559099558741
step: 150, loss: 0.0072820307686924934
step: 160, loss: 0.0004206706362310797
step: 170, loss: 0.0002255165163660422
step: 180, loss: 0.0020177329424768686
step: 190, loss: 0.0006605810485780239
step: 200, loss: 0.0006770096370019019
step: 210, loss: 0.000771129853092134
/usr/local/lib/python3.7/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=0.6696230598669622, f1=0.7096774193548389,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043137066066265106
step: 10, loss: 0.00014978974650148302
step: 20, loss: 0.015593844465911388
step: 30, loss: 0.009117420762777328
step: 40, loss: 0.006928320042788982
step: 50, loss: 0.001461791922338307
step: 60, loss: 0.00033561023883521557
step: 70, loss: 0.028091466054320335
step: 80, loss: 0.0005509564653038979
step: 90, loss: 0.0009417958790436387
step: 100, loss: 0.00024001998826861382
step: 110, loss: 0.00015723895921837538
step: 120, loss: 0.03061138279736042
step: 130, loss: 0.0002742336655501276
step: 140, loss: 0.029529472813010216
step: 150, loss: 0.0008577368571422994
step: 160, loss: 0.0001618647074792534
step: 170, loss: 0.0016584140248596668
step: 180, loss: 0.015033304691314697
step: 190, loss: 0.00031274493085220456
step: 200, loss: 0.000713075278326869
step: 210, loss: 0.000417589268181473
/usr/local/lib/python3.7/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=0.6740088105726872, f1=0.7253218884120172,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015340288518927991

```

```

step: 10, loss: 0.03324170783162117
step: 20, loss: 0.0005685038631781936
step: 30, loss: 0.00011290278052911162
step: 40, loss: 0.00038091407623142004
step: 50, loss: 0.003130493452772498
step: 60, loss: 0.0008634654222987592
step: 70, loss: 0.0004118398064747453
step: 80, loss: 0.0006620979984290898
step: 90, loss: 0.0008161015575751662
step: 100, loss: 0.001582173048518598
step: 110, loss: 0.00021582508634310216
step: 120, loss: 0.03127795085310936
step: 130, loss: 0.01628565788269043
step: 140, loss: 0.003779255785048008
step: 150, loss: 0.0003897678980138153
step: 160, loss: 0.0030111894011497498
step: 170, loss: 0.00012686311674769968
step: 180, loss: 0.00017394903989043087
step: 190, loss: 0.0019159193616360426
step: 200, loss: 0.000511837424710393
step: 210, loss: 0.0004575460043270141
/usr/local/lib/python3.7/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=0.6727272727272726, f1=0.7161572052401747,
best_f1=0.6982922201138521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003675420884974301
step: 10, loss: 7.523484964622185e-05
step: 20, loss: 0.0024847141467034817
step: 30, loss: 0.03710983693599701
step: 40, loss: 0.0017067660810425878
step: 50, loss: 0.00014240706514101475
step: 60, loss: 0.04784611612558365
step: 70, loss: 0.0004139358061365783
step: 80, loss: 0.0003579857002478093
step: 90, loss: 0.00748444302007556
step: 100, loss: 0.0003945365606341511
step: 110, loss: 0.0009291782043874264
step: 120, loss: 0.0002764284727163613
step: 130, loss: 0.00010657058737706393
step: 140, loss: 7.295108662219718e-05
step: 150, loss: 0.0002119808050338179
step: 160, loss: 0.019895382225513458
step: 170, loss: 0.00010101342923007905
step: 180, loss: 0.00014656079292763025

```

```

step: 190, loss: 0.00022986340627539903
step: 200, loss: 0.03390087932348251
step: 210, loss: 0.00026326224906370044
/usr/local/lib/python3.7/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=0.6727272727272726, f1=0.7177242888402624,
best_f1=0.6982922201138521
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 [00:07, 289.14it/s]
load_f1 = 0.7011494252873562
real_f1 = 0.7024952015355086
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 187.73it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

```
[8]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task 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, 429kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 1.24MB/s]

Downloading: 100% 440M/440M [00:09<00:00, 48.3MB/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: 0.5865938663482666
step: 10, loss: 0.35721439123153687
step: 20, loss: 0.2964412569999695
step: 30, loss: 0.44292670488357544
step: 40, loss: 0.41181665658950806
step: 50, loss: 0.311949223279953
step: 60, loss: 0.2551012337207794
step: 70, loss: 0.21731548011302948
step: 80, loss: 0.3177900016307831
step: 90, loss: 0.29860666394233704
step: 100, loss: 0.28974226117134094
step: 110, loss: 0.3945315182209015
step: 120, loss: 0.04810023307800293
step: 130, loss: 0.1360507756471634
step: 140, loss: 0.08535171300172806
step: 150, loss: 0.13212773203849792
step: 160, loss: 0.09497123211622238
step: 170, loss: 0.28415390849113464
step: 180, loss: 0.009999465197324753
step: 190, loss: 0.12672953307628632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6820512820512821, f1=0.7, best_f1=0.7
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20378312468528748
step: 10, loss: 0.05368245020508766
step: 20, loss: 0.06634208559989929
step: 30, loss: 0.14681492745876312
step: 40, loss: 0.15643049776554108
step: 50, loss: 0.11351106315851212
step: 60, loss: 0.15119989216327667
step: 70, loss: 0.190129354596138
step: 80, loss: 0.08833975344896317
step: 90, loss: 0.09790586680173874
step: 100, loss: 0.020582787692546844
step: 110, loss: 0.15655997395515442
step: 120, loss: 0.3209364116191864
step: 130, loss: 0.07332620769739151
step: 140, loss: 0.12367964535951614
step: 150, loss: 0.09017238765954971
step: 160, loss: 0.06122049689292908
step: 170, loss: 0.09558537602424622
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.07505200803279877
step: 190, loss: 0.02157633565366268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7735368956743003, f1=0.7916666666666666,
best_f1=0.7916666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0226306039839983
step: 10, loss: 0.1423933207988739
step: 20, loss: 0.08961259573698044
step: 30, loss: 0.037972431629896164
step: 40, loss: 0.11996380239725113
step: 50, loss: 0.20798973739147186
step: 60, loss: 0.017715388908982277
step: 70, loss: 0.0875021442770958
step: 80, loss: 0.06765350699424744
step: 90, loss: 0.1400030106306076

```



```

step: 100, loss: 0.01304020918905735
step: 110, loss: 0.04121968522667885
step: 120, loss: 0.047711681574583054
step: 130, loss: 0.005427014082670212
step: 140, loss: 0.008845409378409386
step: 150, loss: 0.18343104422092438
step: 160, loss: 0.12368667870759964
step: 170, loss: 0.19637063145637512
step: 180, loss: 0.017565429210662842
step: 190, loss: 0.1708003729581833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8062015503875969, f1=0.825974025974026,
best_f1=0.825974025974026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10474654287099838
step: 10, loss: 0.09606128185987473
step: 20, loss: 0.013571249321103096
step: 30, loss: 0.015620357356965542
step: 40, loss: 0.00988000351935625
step: 50, loss: 0.00784204714000225
step: 60, loss: 0.024426531046628952
step: 70, loss: 0.05216317996382713
step: 80, loss: 0.11449497938156128
step: 90, loss: 0.003046087920665741
step: 100, loss: 0.04922247305512428
step: 110, loss: 0.008979943580925465
step: 120, loss: 0.017953133210539818
step: 130, loss: 0.07366526871919632
step: 140, loss: 0.01559306401759386
step: 150, loss: 0.02733534574508667
step: 160, loss: 0.0023206283804029226
step: 170, loss: 0.038318950682878494
step: 180, loss: 0.0025343061424791813
step: 190, loss: 0.07934823632240295
/usr/local/lib/python3.7/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=0.7988826815642459, f1=0.8087431693989071,
best_f1=0.825974025974026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12636323273181915
step: 10, loss: 0.08140704780817032
step: 20, loss: 0.16230380535125732
step: 30, loss: 0.005725455470383167
step: 40, loss: 0.023203128948807716
step: 50, loss: 0.0471050962805748
step: 60, loss: 0.04604734480381012
step: 70, loss: 0.0037332193460315466
step: 80, loss: 0.017293622717261314
step: 90, loss: 0.006826955825090408
step: 100, loss: 0.0005322694778442383
step: 110, loss: 0.004705913830548525
step: 120, loss: 0.002005487447604537
step: 130, loss: 0.32989364862442017
step: 140, loss: 0.13333146274089813
step: 150, loss: 0.1719728708267212
step: 160, loss: 0.037025872617959976
step: 170, loss: 0.0018094780389219522
step: 180, loss: 0.0053368923254311085
step: 190, loss: 0.041089627891778946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8346883468834688, f1=0.8247978436657681,
best_f1=0.8247978436657681
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003019314259290695
step: 10, loss: 0.17778974771499634
step: 20, loss: 0.0015873659867793322
step: 30, loss: 0.01522390078753233
step: 40, loss: 0.011945206671953201
step: 50, loss: 0.0024050683714449406
step: 60, loss: 0.026474667713046074
step: 70, loss: 0.004806856159120798
step: 80, loss: 0.07287149131298065
step: 90, loss: 0.016783958300948143
step: 100, loss: 0.0011716834269464016
step: 110, loss: 0.059727173298597336
step: 120, loss: 0.024700799956917763
step: 130, loss: 0.002282750094309449

```

```

step: 140, loss: 0.19394385814666748
step: 150, loss: 0.0007098669884726405
step: 160, loss: 0.040253229439258575
step: 170, loss: 0.005522018764168024
step: 180, loss: 0.014391392469406128
step: 190, loss: 0.0036577649880200624
/usr/local/lib/python3.7/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=0.8210526315789474, f1=0.8113695090439276,
best_f1=0.8247978436657681
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006684667896479368
step: 10, loss: 0.0027031642384827137
step: 20, loss: 0.10857662558555603
step: 30, loss: 0.008233821950852871
step: 40, loss: 0.0020224410109221935
step: 50, loss: 0.0008919801912270486
step: 60, loss: 0.0036163809709250927
step: 70, loss: 0.0210188627243042
step: 80, loss: 0.02437932789325714
step: 90, loss: 0.003411638317629695
step: 100, loss: 0.0010443489300087094
step: 110, loss: 0.008814102038741112
step: 120, loss: 0.001933448133058846
step: 130, loss: 0.01503411028534174
step: 140, loss: 0.003874105866998434
step: 150, loss: 0.02220325544476509
step: 160, loss: 0.05284641310572624
step: 170, loss: 0.008907271549105644
step: 180, loss: 0.025871017947793007
step: 190, loss: 0.0018657800974324346
/usr/local/lib/python3.7/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=0.8295165394402036, f1=0.8080808080808081,
best_f1=0.8247978436657681
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008448894135653973
step: 10, loss: 0.018422415480017662
step: 20, loss: 0.002009254414588213
step: 30, loss: 0.007163609843701124
step: 40, loss: 0.0006376924575306475
step: 50, loss: 0.0005012364126741886

```

```

step: 60, loss: 0.0004096931661479175
step: 70, loss: 0.0014584214659407735
step: 80, loss: 0.00017918174853548408
step: 90, loss: 0.008468924090266228
step: 100, loss: 0.1359950602054596
step: 110, loss: 0.0013245171867311
step: 120, loss: 0.0003436552360653877
step: 130, loss: 0.0018944772891700268
step: 140, loss: 0.000625326472800225
step: 150, loss: 0.0010687877656891942
step: 160, loss: 0.009355548769235611
step: 170, loss: 0.001196052529849112
step: 180, loss: 0.02996591106057167
step: 190, loss: 0.20561747252941132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8431876606683805, f1=0.810126582278481,
best_f1=0.810126582278481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007918911869637668
step: 10, loss: 0.005642284639179707
step: 20, loss: 0.002407581778243184
step: 30, loss: 0.004179199226200581
step: 40, loss: 0.03682345896959305
step: 50, loss: 0.000778936839196831
step: 60, loss: 0.001124493544921279
step: 70, loss: 0.0007839045720174909
step: 80, loss: 0.0010753251845017076
step: 90, loss: 0.045654382556676865
step: 100, loss: 0.00849819090217352
step: 110, loss: 0.0005842997343279421
step: 120, loss: 0.00744903739541769
step: 130, loss: 0.0032876215409487486
step: 140, loss: 0.00600839639082551
step: 150, loss: 0.0005831705057062209
step: 160, loss: 0.0003490794915705919
step: 170, loss: 0.005708370357751846
step: 180, loss: 0.018059544265270233
step: 190, loss: 0.06152123212814331
/usr/local/lib/python3.7/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=0.8198433420365535, f1=0.8102564102564102,
best_f1=0.810126582278481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009446402546018362
step: 10, loss: 0.0005370579892769456
step: 20, loss: 0.0015262487577274442
step: 30, loss: 0.0002299731131643057
step: 40, loss: 0.003326851176097989
step: 50, loss: 0.0008404711843468249
step: 60, loss: 0.0007085802499204874
step: 70, loss: 0.0005753687582910061
step: 80, loss: 0.00023555364168714732
step: 90, loss: 0.0005451304605230689
step: 100, loss: 0.0020205979235470295
step: 110, loss: 0.014367996715009212
step: 120, loss: 0.1734209805727005
step: 130, loss: 0.0017852573655545712
step: 140, loss: 0.0018003880977630615
step: 150, loss: 0.000530394958332181
step: 160, loss: 0.0026956319343298674
step: 170, loss: 0.004235722590237856
step: 180, loss: 0.0011681007454171777
step: 190, loss: 0.0007053473382256925
/usr/local/lib/python3.7/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=0.8226221079691518, f1=0.8151898734177214,
best_f1=0.810126582278481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025780368014238775
step: 10, loss: 0.0004555793711915612
step: 20, loss: 0.005670144222676754
step: 30, loss: 0.0006799034890718758
step: 40, loss: 0.00022356152476277202
step: 50, loss: 0.015205263160169125
step: 60, loss: 0.0006867132615298033
step: 70, loss: 0.0002904323046095669
step: 80, loss: 0.0005240521277301013
step: 90, loss: 0.0003830156638287008
step: 100, loss: 0.0002313361328560859
step: 110, loss: 0.00027746977866627276
step: 120, loss: 0.00023271604732144624
step: 130, loss: 0.00039027148159220815

```

```

step: 140, loss: 0.0002719374024309218
step: 150, loss: 0.0004706113541033119
step: 160, loss: 0.00047125876881182194
step: 170, loss: 0.0004245327436365187
step: 180, loss: 0.0002227411896456033
step: 190, loss: 0.00027312073507346213
/usr/local/lib/python3.7/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=0.8268733850129197, f1=0.8020565552699229,
best_f1=0.810126582278481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018950304365716875
step: 10, loss: 0.0014491659821942449
step: 20, loss: 0.00015666171384509653
step: 30, loss: 0.0006233994499780238
step: 40, loss: 0.0004016860912088305
step: 50, loss: 0.00041158648673444986
step: 60, loss: 0.10063621401786804
step: 70, loss: 0.0013917393516749144
step: 80, loss: 0.0006370493792928755
step: 90, loss: 0.00041276513366028666
step: 100, loss: 0.00014241824101191014
step: 110, loss: 0.0004899417399428785
step: 120, loss: 0.00039371338789351285
step: 130, loss: 0.0016390923410654068
step: 140, loss: 0.0006877320702187717
step: 150, loss: 0.0004213425563648343
step: 160, loss: 0.0003000643046107143
step: 170, loss: 0.0018189706606790423
step: 180, loss: 8.81833111634478e-05
step: 190, loss: 0.0003254527982790023
/usr/local/lib/python3.7/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=0.8201058201058201, f1=0.8229166666666667,
best_f1=0.810126582278481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00363214616663754
step: 10, loss: 0.00036801304668188095
step: 20, loss: 0.0003062924079131335
step: 30, loss: 0.00948583148419857
step: 40, loss: 0.00037636718479916453
step: 50, loss: 0.0006543162162415683

```

```

step: 60, loss: 0.0013434337452054024
step: 70, loss: 0.0010481263743713498
step: 80, loss: 0.0013821799075230956
step: 90, loss: 0.0004329865041654557
step: 100, loss: 0.0010722943115979433
step: 110, loss: 0.00015467579942196608
step: 120, loss: 0.00032711392850615084
step: 130, loss: 0.0002189712831750512
step: 140, loss: 0.00021427754836622626
step: 150, loss: 0.0004215278895571828
step: 160, loss: 0.00021813773491885513
step: 170, loss: 0.0006844980525784194
step: 180, loss: 0.00015103677287697792
step: 190, loss: 0.00013869829126633704
/usr/local/lib/python3.7/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=0.8206521739130436, f1=0.8213333333333334,
best_f1=0.810126582278481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007906366372480989
step: 10, loss: 0.00014807426487095654
step: 20, loss: 0.0004436596645973623
step: 30, loss: 0.00018445482419338077
step: 40, loss: 0.0012022866867482662
step: 50, loss: 0.0002846109273377806
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.0004791023675352335
step: 70, loss: 0.00018682435620576143
step: 80, loss: 0.0032490650191903114
step: 90, loss: 0.00039238997851498425
step: 100, loss: 0.001118182553909719
step: 110, loss: 0.001548382337205112
step: 120, loss: 0.0007577294018119574
step: 130, loss: 0.00017774284060578793
step: 140, loss: 0.0007538635400123894
step: 150, loss: 0.0007801384199410677
step: 160, loss: 0.00012904957111459225
step: 170, loss: 0.0003127653617411852
step: 180, loss: 0.00047597583034075797
step: 190, loss: 0.000285242305835709
/usr/local/lib/python3.7/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=0.8275862068965517, f1=0.8306878306878307,
best_f1=0.810126582278481

```

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

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

```
step: 0, loss: 0.0016286657191812992
step: 10, loss: 0.00013152301835361868
step: 20, loss: 0.00021171958360355347
step: 30, loss: 0.00010195164213655517
step: 40, loss: 0.0003220518701709807
step: 50, loss: 0.0003205364919267595
step: 60, loss: 0.00017078546807169914
step: 70, loss: 0.0002317492908332497
step: 80, loss: 0.0006788780447095633
step: 90, loss: 0.0001381609617965296
step: 100, loss: 0.00021397248201537877
step: 110, loss: 0.00020276461145840585
step: 120, loss: 0.00022078024630900472
step: 130, loss: 0.00013308203779160976
step: 140, loss: 0.00122683250810951
step: 150, loss: 0.0009879665449261665
step: 160, loss: 0.00013733989908359945
step: 170, loss: 0.00019833806436508894
step: 180, loss: 0.0002903647255152464
step: 190, loss: 0.0001885163801489398
```

```
/usr/local/lib/python3.7/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=0.830379746835443, f1=0.8193384223918575,
best_f1=0.810126582278481
```

```
mv: cannot stat '*_dev.pt': No such file or directory
```

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

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

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

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

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

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



```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:09, 226.38it/s]
load_f1 = 0.8393782383419688
real_f1 = 0.8372093023255814
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, 245.04it/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: 0.6141785383224487  
step: 10, loss: 0.34808823466300964  
step: 20, loss: 0.2942225933074951  
step: 30, loss: 0.3712674081325531  
step: 40, loss: 0.2514421343803406  
step: 50, loss: 0.18770891427993774  
step: 60, loss: 0.33784908056259155
```

```

step: 70, loss: 0.2784614861011505
step: 80, loss: 0.2142651528120041
step: 90, loss: 0.1924521028995514
step: 100, loss: 0.2680931091308594
step: 110, loss: 0.20302094519138336
step: 120, loss: 0.04834022745490074
step: 130, loss: 0.015080355107784271
step: 140, loss: 0.09626909345388412
step: 150, loss: 0.21519574522972107
step: 160, loss: 0.15782834589481354
step: 170, loss: 0.13239680230617523
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7971014492753623, f1=0.7659574468085106,
best_f1=0.7659574468085106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2094152569770813
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.33066219091415405
step: 20, loss: 0.08120573312044144
step: 30, loss: 0.1228925958275795
step: 40, loss: 0.08757514506578445
step: 50, loss: 0.050237081944942474
step: 60, loss: 0.046962033957242966
step: 70, loss: 0.07495676726102829
step: 80, loss: 0.06889521330595016
step: 90, loss: 0.11235977709293365
step: 100, loss: 0.0777464210987091
step: 110, loss: 0.024273652583360672
step: 120, loss: 0.07704359292984009
step: 130, loss: 0.09557696431875229
step: 140, loss: 0.2954353988170624
step: 150, loss: 0.16221649944782257
step: 160, loss: 0.12853367626667023
step: 170, loss: 0.028911802917718887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7990074441687345, f1=0.8067632850241547,
best_f1=0.8067632850241547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08814091235399246
step: 10, loss: 0.06022956222295761
step: 20, loss: 0.048405032604932785
step: 30, loss: 0.12821519374847412
step: 40, loss: 0.06426823884248734
step: 50, loss: 0.19252146780490875
step: 60, loss: 0.11206137388944626
step: 70, loss: 0.033679261803627014
step: 80, loss: 0.061946846544742584
step: 90, loss: 0.0680948942899704
step: 100, loss: 0.016300790011882782
step: 110, loss: 0.09495531022548676
step: 120, loss: 0.05309325456619263
step: 130, loss: 0.13612164556980133
step: 140, loss: 0.04989247769117355
step: 150, loss: 0.010654973797500134
step: 160, loss: 0.0063363634981215
step: 170, loss: 0.07682839035987854
/usr/local/lib/python3.7/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=0.7889908256880733, f1=0.7789934354485777,
best_f1=0.8067632850241547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011631855741143227
step: 10, loss: 0.07592564076185226
step: 20, loss: 0.008249210193753242
step: 30, loss: 0.031646981835365295
step: 40, loss: 0.004456623923033476
step: 50, loss: 0.027266310527920723
step: 60, loss: 0.06829079985618591
step: 70, loss: 0.0072541069239377975
step: 80, loss: 0.1738440841436386
step: 90, loss: 0.03396693244576454
step: 100, loss: 0.15994225442409515
step: 110, loss: 0.13145671784877777
step: 120, loss: 0.012924136593937874
step: 130, loss: 0.06236695498228073
step: 140, loss: 0.04251232370734215
step: 150, loss: 0.09202674776315689

```

```

step: 160, loss: 0.2902693450450897
step: 170, loss: 0.013960187323391438
/usr/local/lib/python3.7/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=0.7990074441687345, f1=0.8028846153846153,
best_f1=0.8067632850241547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028083214536309242
step: 10, loss: 0.017199723049998283
step: 20, loss: 0.003745194524526596
step: 30, loss: 0.14407622814178467
step: 40, loss: 0.01823442056775093
step: 50, loss: 0.00422405730932951
step: 60, loss: 0.0064247683621943
step: 70, loss: 0.2638779282569885
step: 80, loss: 0.04111558571457863
step: 90, loss: 0.062424659729003906
step: 100, loss: 0.031672678887844086
step: 110, loss: 0.08711277693510056
step: 120, loss: 0.014640157110989094
step: 130, loss: 0.06761255860328674
step: 140, loss: 0.005893700756132603
step: 150, loss: 0.025882825255393982
step: 160, loss: 0.01137203723192215
step: 170, loss: 0.01059684157371521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8082901554404144, f1=0.8388746803069055,
best_f1=0.8388746803069055
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002171242143958807
step: 10, loss: 0.010744954459369183
step: 20, loss: 0.008577787317335606
step: 30, loss: 0.018602143973112106
step: 40, loss: 0.009055178612470627
step: 50, loss: 0.17318983376026154
step: 60, loss: 0.021032098680734634
step: 70, loss: 0.04691273719072342

```

```

step: 80, loss: 0.03016536496579647
step: 90, loss: 0.07232643663883209
step: 100, loss: 0.016946902498602867
step: 110, loss: 0.0028374199755489826
step: 120, loss: 0.021087946370244026
step: 130, loss: 0.009195128455758095
step: 140, loss: 0.01538023166358471
step: 150, loss: 0.0626906231045723
step: 160, loss: 0.07139856368303299
step: 170, loss: 0.007764921523630619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.836734693877551, f1=0.8312342569269522,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015462550800293684
step: 10, loss: 0.00729077123105526
step: 20, loss: 0.004577529150992632
step: 30, loss: 0.016687870025634766
step: 40, loss: 0.0029744247440248728
step: 50, loss: 0.010566258803009987
step: 60, loss: 0.0002596874546725303
step: 70, loss: 0.0004162908880971372
step: 80, loss: 0.10230965912342072
step: 90, loss: 0.0005693449638783932
step: 100, loss: 0.0010425288928672671
step: 110, loss: 0.012303706258535385
step: 120, loss: 0.004661035258322954
step: 130, loss: 0.11799363791942596
step: 140, loss: 0.002096671611070633
step: 150, loss: 0.0009049968793988228
step: 160, loss: 0.00044943945249542594
step: 170, loss: 0.05899426341056824
/usr/local/lib/python3.7/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=0.8203753351206435, f1=0.8226221079691517,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.02712869457900524
step: 10, loss: 0.001345787663012743
step: 20, loss: 0.00034427776699885726
step: 30, loss: 0.027135778218507767
step: 40, loss: 0.00025201222160831094
step: 50, loss: 0.002524537965655327
step: 60, loss: 0.0018160805338993669
step: 70, loss: 0.005041064694523811
step: 80, loss: 0.031484488397836685
step: 90, loss: 0.001368184108287096
step: 100, loss: 0.027267025783658028
step: 110, loss: 0.06359675526618958
step: 120, loss: 0.00036858435487374663
step: 130, loss: 0.006867910269647837
step: 140, loss: 0.002411292167380452
step: 150, loss: 0.004099732264876366
step: 160, loss: 0.03186170384287834
step: 170, loss: 0.0022839494049549103
/usr/local/lib/python3.7/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=0.8350515463917526, f1=0.8350000000000001,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007024262449704111
step: 10, loss: 0.006583303678780794
step: 20, loss: 0.1727575957775116
step: 30, loss: 0.0030511498916894197
step: 40, loss: 0.005670074373483658
step: 50, loss: 0.00029936485225334764
step: 60, loss: 0.007745716720819473
step: 70, loss: 0.01396129745990038
step: 80, loss: 0.0008066685986705124
step: 90, loss: 0.010084333829581738
step: 100, loss: 0.006750875618308783
step: 110, loss: 0.0001998868683585897
step: 120, loss: 0.006112896837294102
step: 130, loss: 0.06232353672385216
step: 140, loss: 0.0011533605866134167
step: 150, loss: 0.0008218840230256319
step: 160, loss: 0.0016411736141890287
step: 170, loss: 0.033422261476516724
/usr/local/lib/python3.7/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=0.8337468982630273, f1=0.8454106280193237,

```

```

best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04834124073386192
step: 10, loss: 0.0017135434318333864
step: 20, loss: 0.07761434465646744
step: 30, loss: 0.04876530542969704
step: 40, loss: 0.00043815464596264064
step: 50, loss: 0.08030983060598373
step: 60, loss: 0.00024356551875825971
step: 70, loss: 0.0163719542324543
step: 80, loss: 0.008161483332514763
step: 90, loss: 0.009911163710057735
step: 100, loss: 0.0002746792451944202
step: 110, loss: 0.005130785051733255
step: 120, loss: 0.00037879470619373024
step: 130, loss: 0.0012479068245738745
step: 140, loss: 0.001857961993664503
step: 150, loss: 0.0003149382537230849
step: 160, loss: 0.0008344536763615906
step: 170, loss: 0.00023034815967548639
/usr/local/lib/python3.7/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=0.8195121951219512, f1=0.8221709006928406,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05635756254196167
step: 10, loss: 0.008411690592765808
step: 20, loss: 0.0001270071225007996
step: 30, loss: 0.00012139136379119009
step: 40, loss: 0.0005416994681581855
step: 50, loss: 0.004068948794156313
step: 60, loss: 0.00041909742867574096
step: 70, loss: 8.586221520090476e-05
step: 80, loss: 6.854190723970532e-05
step: 90, loss: 9.518786100670695e-05
step: 100, loss: 0.004527730867266655
step: 110, loss: 0.005393756553530693
step: 120, loss: 0.00013861515617463738
step: 130, loss: 8.816735498839989e-05
step: 140, loss: 0.00014180516882333905
step: 150, loss: 0.0005619522416964173
step: 160, loss: 0.003519026329740882
step: 170, loss: 0.0009464927716180682

```



```

/usr/local/lib/python3.7/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=0.8329297820823245, f1=0.8287037037037036,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005843703169375658
step: 10, loss: 0.0002672444679774344
step: 20, loss: 0.0001448667753720656
step: 30, loss: 0.00016706525639165193
step: 40, loss: 0.00020600615243893117
step: 50, loss: 8.425271516898647e-05
step: 60, loss: 0.00016781300655566156
step: 70, loss: 0.10372979938983917
step: 80, loss: 0.0016899722395464778
step: 90, loss: 0.00028419416048564017
step: 100, loss: 0.000776418368332088
step: 110, loss: 0.0001797753357095644
step: 120, loss: 0.015365502797067165
step: 130, loss: 0.013144818134605885
step: 140, loss: 0.002915431512519717
step: 150, loss: 0.0030286312103271484
step: 160, loss: 0.00010430812108097598
step: 170, loss: 0.00012329402670729905
/usr/local/lib/python3.7/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=0.7989556135770236, f1=0.8215158924205379,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012356976047158241
step: 10, loss: 0.0003056526184082031
step: 20, loss: 0.0010362438624724746
step: 30, loss: 0.00013628709712065756
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.00021063379244878888
step: 50, loss: 0.00011051573528675362
step: 60, loss: 0.013029390014708042
step: 70, loss: 0.0011005040723830462
step: 80, loss: 0.0012537278234958649
step: 90, loss: 0.00019735240493901074
step: 100, loss: 0.0006938619771972299
step: 110, loss: 6.0879287048010156e-05
step: 120, loss: 0.012126176618039608

```

```

step: 130, loss: 0.02759888581931591
step: 140, loss: 0.00047643418656662107
step: 150, loss: 0.0038753817789256573
step: 160, loss: 0.010586101561784744
step: 170, loss: 0.0003250096342526376
/usr/local/lib/python3.7/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=0.8219895287958117, f1=0.838709677419355,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003908707294613123
step: 10, loss: 0.0030292808078229427
step: 20, loss: 0.00041882990626618266
step: 30, loss: 0.0016114400932565331
step: 40, loss: 7.077869668137282e-05
step: 50, loss: 0.00012154633441241458
step: 60, loss: 4.7674679080955684e-05
step: 70, loss: 0.00019487716781441122
step: 80, loss: 0.06719578057527542
step: 90, loss: 7.336798444157466e-05
step: 100, loss: 7.710707723163068e-05
step: 110, loss: 0.00010131150338565931
step: 120, loss: 0.10144273936748505
step: 130, loss: 0.0010684998705983162
step: 140, loss: 0.014861319214105606
step: 150, loss: 6.60455843899399e-05
step: 160, loss: 4.2549210775177926e-05
step: 170, loss: 0.00025092627038247883
/usr/local/lib/python3.7/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=0.8250652741514359, f1=0.838709677419355,
best_f1=0.8312342569269522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011429552250774577
step: 10, loss: 0.001644724397920072
step: 20, loss: 0.0050892638973891735
step: 30, loss: 0.0014784560771659017
step: 40, loss: 0.0017389252316206694
step: 50, loss: 0.0002149172651115805
step: 60, loss: 0.06664387136697769
step: 70, loss: 6.994597060838714e-05
step: 80, loss: 0.023477977141737938

```

```

step: 90, loss: 0.005588054191321135
step: 100, loss: 0.0004936035838909447
step: 110, loss: 8.03728835307993e-05
step: 120, loss: 0.00017428775026928633
step: 130, loss: 0.003671855665743351
step: 140, loss: 0.00010366681817686185
step: 150, loss: 0.00015986144717317075
step: 160, loss: 0.0001883225340861827
step: 170, loss: 0.00013468399993143976
/usr/local/lib/python3.7/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=0.828496042216359, f1=0.8535353535353535,
best_f1=0.8312342569269522
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, 257.12it/s]
load_f1 = 0.832080200501253
real_f1 = 0.8341708542713568
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 234.16it/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: 0.6284317970275879
step: 10, loss: 0.6186806559562683
step: 20, loss: 0.4216397702693939
step: 30, loss: 0.1498863697052002
step: 40, loss: 0.19390717148780823
step: 50, loss: 0.10719646513462067
step: 60, loss: 0.12753665447235107
step: 70, loss: 0.06160859018564224
step: 80, loss: 0.032266415655612946
step: 90, loss: 0.09256072342395782
step: 100, loss: 0.014699172228574753
step: 110, loss: 0.10654760152101517
step: 120, loss: 0.012693806551396847
step: 130, loss: 0.008816308341920376
step: 140, loss: 0.003302360186353326
step: 150, loss: 0.03924860432744026
step: 160, loss: 0.12204895168542862
step: 170, loss: 0.13394580781459808
step: 180, loss: 0.07284615188837051
step: 190, loss: 0.05298718810081482
step: 200, loss: 0.029445169493556023
step: 210, loss: 0.00867969449609518

```

```

step: 220, loss: 0.019013892859220505
step: 230, loss: 0.006957183592021465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9754464285714286, f1=0.9673790776152981,
best_f1=0.9673790776152981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009334543719887733
step: 10, loss: 0.0020616084802895784
step: 20, loss: 0.096721351146698
step: 30, loss: 0.18979553878307343
step: 40, loss: 0.11306869238615036
step: 50, loss: 0.004374097101390362
step: 60, loss: 0.0044717383570969105
step: 70, loss: 0.03719629719853401
step: 80, loss: 0.00438608368858695
step: 90, loss: 0.006893565878272057
step: 100, loss: 0.04877960681915283
step: 110, loss: 0.1509529948234558
step: 120, loss: 0.0941324383020401
step: 130, loss: 0.09666624665260315
step: 140, loss: 0.003106793388724327
step: 150, loss: 0.09336026012897491
step: 160, loss: 0.033927805721759796
step: 170, loss: 0.0016696159727871418
step: 180, loss: 0.004250967409461737
step: 190, loss: 0.0015879245474934578
step: 200, loss: 0.001420601736754179
step: 210, loss: 0.0006338771781884134
step: 220, loss: 0.20677529275417328
step: 230, loss: 0.015700258314609528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831649831649831,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006311275064945221
step: 10, loss: 0.0029532138723880053
step: 20, loss: 0.0014860244700685143
step: 30, loss: 0.0037187752313911915
step: 40, loss: 0.05385829508304596
step: 50, loss: 0.008325867354869843
step: 60, loss: 0.008610044606029987
step: 70, loss: 0.003281879471614957
step: 80, loss: 0.00261782668530941
step: 90, loss: 0.030133826658129692
step: 100, loss: 0.0019266590243205428
step: 110, loss: 0.0013435985893011093
step: 120, loss: 0.010837594978511333
step: 130, loss: 0.0005181044107303023
step: 140, loss: 0.013213614001870155
step: 150, loss: 0.08225667476654053
step: 160, loss: 0.13785968720912933
step: 170, loss: 0.037647295743227005
step: 180, loss: 0.011063250713050365
step: 190, loss: 0.01037878729403019
step: 200, loss: 0.037328023463487625
step: 210, loss: 0.2025538682937622
step: 220, loss: 0.001894910354167223
step: 230, loss: 0.0008491053013131022
/usr/local/lib/python3.7/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=0.978675645342312, f1=0.9751693002257337,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007643329910933971
step: 10, loss: 0.009239769540727139
step: 20, loss: 0.0005952953360974789
step: 30, loss: 0.0008338623447343707
step: 40, loss: 0.005378290545195341
step: 50, loss: 0.0006294230697676539
step: 60, loss: 0.0024101226590573788
step: 70, loss: 0.0009335969807580113
step: 80, loss: 0.0024907588958740234
step: 90, loss: 0.005397876724600792
step: 100, loss: 0.00037693374906666577
step: 110, loss: 0.0004287678166292608
step: 120, loss: 0.01142097357660532
step: 130, loss: 0.0010361992754042149

```

```

step: 140, loss: 0.00030783412512391806
step: 150, loss: 0.10677451640367508
step: 160, loss: 0.0025749285705387592
step: 170, loss: 0.007921753451228142
step: 180, loss: 0.0008836453198455274
step: 190, loss: 0.0018399265827611089
step: 200, loss: 0.000576353573706001
step: 210, loss: 0.15433022379875183
step: 220, loss: 0.0006700336816720665
step: 230, loss: 0.0032940448727458715
/usr/local/lib/python3.7/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=0.978675645342312, f1=0.9785794813979707,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013849546667188406
step: 10, loss: 0.0011175936087965965
step: 20, loss: 0.0023734583519399166
step: 30, loss: 0.0005391753511503339
step: 40, loss: 0.000610777351539582
step: 50, loss: 0.00666741793975234
step: 60, loss: 0.029591232538223267
step: 70, loss: 0.00043800115236081183
step: 80, loss: 0.0003981929621659219
step: 90, loss: 0.0005469652824103832
step: 100, loss: 0.00025771671789698303
step: 110, loss: 0.0007527945563197136
step: 120, loss: 0.0005855869385413826
step: 130, loss: 0.0007391809485852718
step: 140, loss: 0.0025509975384920835
step: 150, loss: 0.0024550689850002527
step: 160, loss: 0.00037347988109104335
step: 170, loss: 0.00625831400975585
step: 180, loss: 0.00023242761380970478
step: 190, loss: 0.057595767080783844
step: 200, loss: 0.0003787697060033679
step: 210, loss: 0.0004011720302514732
step: 220, loss: 0.0006720382370986044
step: 230, loss: 0.0002891462645493448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831649831649831, f1=0.9787234042553192,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037716547958552837
step: 10, loss: 0.00022389396326616406
step: 20, loss: 0.00037434458499774337
step: 30, loss: 0.00016113901801872998
step: 40, loss: 0.00017741993360687047
step: 50, loss: 0.0015137805603444576
step: 60, loss: 8.442738180747256e-05
step: 70, loss: 0.00015270670701283962
step: 80, loss: 0.00022811770031694323
step: 90, loss: 0.00014685180212836713
step: 100, loss: 0.04897652193903923
step: 110, loss: 0.00020982253772672266
step: 120, loss: 0.00044338166480883956
step: 130, loss: 0.0009254192700609565
step: 140, loss: 0.00011120915587525815
step: 150, loss: 0.035721685737371445
step: 160, loss: 0.0005469738971441984
step: 170, loss: 0.00015392441127914935
step: 180, loss: 0.013375286012887955
step: 190, loss: 0.03378608822822571
step: 200, loss: 0.0002490413608029485
step: 210, loss: 0.0011889786692336202
step: 220, loss: 0.0001582536642672494
step: 230, loss: 0.05212794989347458
/usr/local/lib/python3.7/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=0.9820224719101124, f1=0.9785794813979707,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028661221731454134
step: 10, loss: 9.345744183519855e-05
step: 20, loss: 0.000677076110150665
step: 30, loss: 0.0011421446688473225
step: 40, loss: 0.00022416516731027514
step: 50, loss: 0.00010903003567364067
step: 60, loss: 0.0010711266659200191
step: 70, loss: 0.029343796893954277
step: 80, loss: 0.002289881929755211
step: 90, loss: 0.0006738107767887414

```

```

step: 100, loss: 0.0003062354517169297
step: 110, loss: 0.001065892749466002
step: 120, loss: 0.00022469469695352018
step: 130, loss: 5.85899870202411e-05
step: 140, loss: 0.00015550496755167842
step: 150, loss: 0.0014734016731381416
step: 160, loss: 0.07633619755506516
step: 170, loss: 0.03990896791219711
step: 180, loss: 0.00021850873599760234
step: 190, loss: 0.0017744343494996428
step: 200, loss: 0.040689125657081604
step: 210, loss: 0.0003449766372796148
step: 220, loss: 0.0003000311553478241
step: 230, loss: 0.0006719101220369339
/usr/local/lib/python3.7/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=0.9778761061946903, f1=0.9702315325248071,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.186624083667994e-05
step: 10, loss: 0.0003653363964986056
step: 20, loss: 9.200812928611413e-05
step: 30, loss: 7.850499241612852e-05
step: 40, loss: 0.0011395136825740337
step: 50, loss: 0.00021697963529732078
step: 60, loss: 0.00011034969793399796
step: 70, loss: 0.00024694454623386264
step: 80, loss: 0.00021233866573311388
step: 90, loss: 0.00014427486166823655
step: 100, loss: 0.00014088496391195804
step: 110, loss: 0.04667180776596069
step: 120, loss: 0.007502846885472536
step: 130, loss: 0.0002064080908894539
step: 140, loss: 7.707384065724909e-05
step: 150, loss: 7.940520299598575e-05
step: 160, loss: 0.00022354326210916042
step: 170, loss: 0.00015110969252418727
step: 180, loss: 0.0009451029472984374
step: 190, loss: 0.00016969582065939903
step: 200, loss: 0.00027309192228130996
step: 210, loss: 0.00025413630646653473
step: 220, loss: 0.020426126196980476
step: 230, loss: 0.00015540706226602197
/usr/local/lib/python3.7/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=0.9831649831649831, f1=0.9853768278965129,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.710335972253233e-05
step: 10, loss: 0.0003953944251406938
step: 20, loss: 0.0010787993669509888
step: 30, loss: 0.0002165395562769845
step: 40, loss: 0.015790512785315514
step: 50, loss: 5.164439426152967e-05
step: 60, loss: 8.070204057730734e-05
step: 70, loss: 0.0002557438565418124
step: 80, loss: 8.706087101018056e-05
step: 90, loss: 0.0001499130594311282
step: 100, loss: 0.000317699508741498
step: 110, loss: 5.9222908021183684e-05
step: 120, loss: 4.9689377192407846e-05
step: 130, loss: 5.4343166993930936e-05
step: 140, loss: 5.120767673361115e-05
step: 150, loss: 0.0007337993010878563
step: 160, loss: 0.00016328859783243388
step: 170, loss: 8.109195914585143e-05
step: 180, loss: 0.0001715891412459314
step: 190, loss: 4.978542710887268e-05
step: 200, loss: 0.014419769868254662
step: 210, loss: 5.439138112706132e-05
step: 220, loss: 4.845254807150923e-05
step: 230, loss: 0.0002501210256014019
/usr/local/lib/python3.7/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=0.9811320754716982, f1=0.9766925638179801,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.87950565002393e-05
step: 10, loss: 0.0003760012914426625
step: 20, loss: 5.369149221223779e-05
step: 30, loss: 4.4343927584122866e-05
step: 40, loss: 0.0005671848193742335
step: 50, loss: 4.346051355241798e-05
step: 60, loss: 5.034191417507827e-05
step: 70, loss: 7.573521725134924e-05
step: 80, loss: 5.4038184316596016e-05
step: 90, loss: 0.00021616305457428098

```

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 100, loss: 9.176047024084255e-05
step: 110, loss: 0.004298769403249025
step: 120, loss: 0.0006117006996646523
step: 130, loss: 0.00013440031034406275
step: 140, loss: 0.041553787887096405
step: 150, loss: 0.00839389581233263
step: 160, loss: 7.161207759054378e-05
step: 170, loss: 7.632599590579048e-05
step: 180, loss: 9.962018521036953e-05
step: 190, loss: 0.003017979674041271
step: 200, loss: 7.520602957811207e-05
step: 210, loss: 0.00022219338279683143
step: 220, loss: 8.089260518318042e-05
step: 230, loss: 0.0004298409912735224
/usr/local/lib/python3.7/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=0.9822222222222222, f1=0.9799554565701558,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.123469549696892e-05
step: 10, loss: 0.00016809608496259898
step: 20, loss: 0.00025671871844679117
step: 30, loss: 0.0017950108740478754
step: 40, loss: 9.644097735872492e-05
step: 50, loss: 0.00015150575200095773
step: 60, loss: 0.00046891524107195437
step: 70, loss: 7.714171806583181e-05
step: 80, loss: 5.5749242164893076e-05
step: 90, loss: 6.217592454049736e-05
step: 100, loss: 6.987413507886231e-05
step: 110, loss: 0.04850730299949646
step: 120, loss: 0.00010913180449279025
step: 130, loss: 3.0296732802526094e-05
step: 140, loss: 7.315123366424814e-05
step: 150, loss: 0.013776695355772972
step: 160, loss: 4.0509556129109114e-05
step: 170, loss: 0.024812187999486923
step: 180, loss: 0.00010858707537408918
step: 190, loss: 4.631311821867712e-05
step: 200, loss: 8.946924208430573e-05
step: 210, loss: 4.107279164600186e-05
step: 220, loss: 5.9211273764958605e-05
step: 230, loss: 4.9579339247429743e-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=0.9842696629213483, f1=0.9808342728297633,
best_f1=0.9808342728297633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.425017297966406e-05
step: 10, loss: 4.3836189433932304e-05
step: 20, loss: 3.732950062840246e-05
step: 30, loss: 7.798073056619614e-05
step: 40, loss: 6.730541645083576e-05
step: 50, loss: 0.00019693098147399724
step: 60, loss: 0.0005091662751510739
step: 70, loss: 0.00013930271961726248
step: 80, loss: 0.00012584247451741248
step: 90, loss: 4.699442069977522e-05
step: 100, loss: 3.23119675158523e-05
step: 110, loss: 3.982119233114645e-05
step: 120, loss: 0.0002273361897096038
step: 130, loss: 3.5854714951710775e-05
step: 140, loss: 9.168574615614489e-05
step: 150, loss: 8.469026943203062e-05
step: 160, loss: 7.553498289780691e-05
step: 170, loss: 4.329705552663654e-05
step: 180, loss: 0.00012089342635590583
step: 190, loss: 6.008851414662786e-05
step: 200, loss: 2.411690729786642e-05
step: 210, loss: 3.847676271107048e-05
step: 220, loss: 4.1627296013757586e-05
step: 230, loss: 0.044306837022304535
/usr/local/lib/python3.7/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=0.9819819819819819,
best_f1=0.9808342728297633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.35578398942016e-05
step: 10, loss: 0.00020932580810040236
step: 20, loss: 5.2802421123487875e-05
step: 30, loss: 6.054896584828384e-05
step: 40, loss: 4.6011056838324293e-05

```

```

step: 50, loss: 5.7860012020682916e-05
step: 60, loss: 0.00021728395950049162
step: 70, loss: 6.366738671204075e-05
step: 80, loss: 4.238382098264992e-05
step: 90, loss: 7.067358092172071e-05
step: 100, loss: 2.6508500013733283e-05
step: 110, loss: 0.0014811985893175006
step: 120, loss: 0.03228044882416725
step: 130, loss: 4.5513221266446635e-05
step: 140, loss: 6.42402665107511e-05
step: 150, loss: 0.0001401876361342147
step: 160, loss: 0.0013822823530063033
step: 170, loss: 5.8554625866236165e-05
step: 180, loss: 3.981733607361093e-05
step: 190, loss: 3.1179097277345136e-05
step: 200, loss: 0.00020071548351552337
step: 210, loss: 2.2142678062664345e-05
step: 220, loss: 3.867359919240698e-05
step: 230, loss: 3.4285549190826714e-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=0.9854423292273236, f1=0.9810055865921787,
best_f1=0.9810055865921787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.570327918278053e-05
step: 10, loss: 7.1388429205399e-05
step: 20, loss: 3.6964614992029965e-05
step: 30, loss: 4.626360168913379e-05
step: 40, loss: 0.014515381306409836
step: 50, loss: 5.037233859184198e-05
step: 60, loss: 3.485970592009835e-05
step: 70, loss: 4.1921761294361204e-05
step: 80, loss: 5.496219091583043e-05
step: 90, loss: 0.00043105825898237526
step: 100, loss: 3.317614027764648e-05
step: 110, loss: 0.0002656492288224399
step: 120, loss: 2.7208501705899835e-05
step: 130, loss: 3.3913966035470366e-05
step: 140, loss: 9.10202506929636e-05
step: 150, loss: 3.5812223359243944e-05
step: 160, loss: 0.0026159649714827538

```

```

step: 170, loss: 2.9765982617391273e-05
step: 180, loss: 7.941910735098645e-05
step: 190, loss: 4.8062480345834047e-05
step: 200, loss: 3.692211976158433e-05
step: 210, loss: 4.077119956491515e-05
step: 220, loss: 5.7961809943662956e-05
step: 230, loss: 0.002383299171924591
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9866071428571428, f1=0.9810055865921787,
best_f1=0.9810055865921787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.4809769456624053e-05
step: 10, loss: 1.8514387193135917e-05
step: 20, loss: 3.6111650842940435e-05
step: 30, loss: 9.14167394512333e-05
step: 40, loss: 5.996307299938053e-05
step: 50, loss: 0.019226472824811935
step: 60, loss: 4.5032677007839084e-05
step: 70, loss: 5.552321090362966e-05
step: 80, loss: 4.9775615480029956e-05
step: 90, loss: 0.0001060051508829929
step: 100, loss: 4.96689863211941e-05
step: 110, loss: 2.8203317924635485e-05
step: 120, loss: 5.6234715884784237e-05
step: 130, loss: 9.87779931165278e-05
step: 140, loss: 3.0911582143744454e-05
step: 150, loss: 0.0013483378570526838
step: 160, loss: 4.286287003196776e-05
step: 170, loss: 2.461979784129653e-05
step: 180, loss: 6.900565495016053e-05
step: 190, loss: 5.8602610806701705e-05
step: 200, loss: 3.695670238812454e-05
step: 210, loss: 7.632553024450317e-05
step: 220, loss: 6.326541188172996e-05
step: 230, loss: 3.2878186175366864e-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=0.9866071428571428, f1=0.9799107142857142,
best_f1=0.9810055865921787

```

```
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 [00:14, 174.71it/s]
load_f1 = 0.9854096520763187
real_f1 = 0.9854096520763187
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:18, 233.26it/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: 0.6416224241256714
step: 10, loss: 0.64085453748703
step: 20, loss: 0.5481995344161987
step: 30, loss: 0.18813057243824005
step: 40, loss: 0.1888556182384491
step: 50, loss: 0.21807603538036346
step: 60, loss: 0.11436399072408676
step: 70, loss: 0.08818681538105011
step: 80, loss: 0.0397803857922554
step: 90, loss: 0.09106141328811646
step: 100, loss: 0.010645519942045212
step: 110, loss: 0.05560154840350151
step: 120, loss: 0.13908636569976807
step: 130, loss: 0.13636676967144012
step: 140, loss: 0.06937200576066971
step: 150, loss: 0.024542929604649544
step: 160, loss: 0.025323623791337013
step: 170, loss: 0.18641650676727295
step: 180, loss: 0.028645461425185204
step: 190, loss: 0.021369481459259987
step: 200, loss: 0.16856235265731812
step: 210, loss: 0.06129893660545349
step: 220, loss: 0.10693042725324631
step: 230, loss: 0.1689041256904602
step: 240, loss: 0.030749717727303505
step: 250, loss: 0.023039180785417557
step: 260, loss: 0.15937048196792603
step: 270, loss: 0.01816541515290737
step: 280, loss: 0.05388038605451584
step: 290, loss: 0.037125371396541595
step: 300, loss: 0.012587044388055801
step: 310, loss: 0.10295433551073074
step: 320, loss: 0.11812062561511993
step: 330, loss: 0.024478614330291748
step: 340, loss: 0.021254681050777435
step: 350, loss: 0.02713429182767868
step: 360, loss: 0.02637099102139473
```

```

step: 370, loss: 0.11892210692167282
step: 380, loss: 0.04819725081324577
step: 390, loss: 0.18044202029705048
step: 400, loss: 0.26377052068710327
step: 410, loss: 0.04102417081594467
step: 420, loss: 0.15831118822097778
step: 430, loss: 0.14700348675251007
step: 440, loss: 0.03053070418536663
step: 450, loss: 0.005807857029139996
step: 460, loss: 0.017038123682141304
step: 470, loss: 0.12599533796310425
step: 480, loss: 0.05453910306096077
step: 490, loss: 0.05598609149456024
step: 500, loss: 0.041714124381542206
step: 510, loss: 0.027664663270115852
step: 520, loss: 0.06197081506252289
step: 530, loss: 0.004247317090630531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9388888888888889, f1=0.9364269141531322,
best_f1=0.9364269141531322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06782130151987076
step: 10, loss: 0.012063371948897839
step: 20, loss: 0.050775304436683655
step: 30, loss: 0.017968617379665375
step: 40, loss: 0.08306828886270523
step: 50, loss: 0.10619442164897919
step: 60, loss: 0.011668301187455654
step: 70, loss: 0.017029179260134697
step: 80, loss: 0.08635939657688141
step: 90, loss: 0.06012513116002083
step: 100, loss: 0.07494135946035385
step: 110, loss: 0.023155804723501205
step: 120, loss: 0.10305379331111908
step: 130, loss: 0.2848453223705292
step: 140, loss: 0.016170838847756386
step: 150, loss: 0.05225815251469612
step: 160, loss: 0.015289987437427044
step: 170, loss: 0.07202820479869843
step: 180, loss: 0.05470416322350502

```

```

step: 190, loss: 0.046318184584379196
step: 200, loss: 0.00776873342692852
step: 210, loss: 0.02873651310801506
step: 220, loss: 0.041434288024902344
step: 230, loss: 0.005551027599722147
step: 240, loss: 0.01728215254843235
step: 250, loss: 0.053105782717466354
step: 260, loss: 0.008511489257216454
step: 270, loss: 0.12451490759849548
step: 280, loss: 0.013765980489552021
step: 290, loss: 0.023588908836245537
step: 300, loss: 0.13991573452949524
step: 310, loss: 0.014217648655176163
step: 320, loss: 0.14038193225860596
step: 330, loss: 0.04434928297996521
step: 340, loss: 0.021681953221559525
step: 350, loss: 0.0008600680739618838
step: 360, loss: 0.08365516364574432
step: 370, loss: 0.11143246293067932
step: 380, loss: 0.05060967803001404
step: 390, loss: 0.040853239595890045
step: 400, loss: 0.028226574882864952
step: 410, loss: 0.010565420612692833
step: 420, loss: 0.018127908930182457
step: 430, loss: 0.01552165113389492
step: 440, loss: 0.16127771139144897
step: 450, loss: 0.015904879197478294
step: 460, loss: 0.012724214233458042
step: 470, loss: 0.02243778109550476
step: 480, loss: 0.20136409997940063
step: 490, loss: 0.012442514300346375
step: 500, loss: 0.22564467787742615
step: 510, loss: 0.01505922619253397
step: 520, loss: 0.0712563768029213
step: 530, loss: 0.05710840970277786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9475638051044084, f1=0.9400836042731072,
best_f1=0.9400836042731072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03379345312714577

```

step: 10, loss: 0.030557651072740555
step: 20, loss: 0.022860299795866013
step: 30, loss: 0.10595763474702835
step: 40, loss: 0.011903266422450542
step: 50, loss: 0.03785410150885582
step: 60, loss: 0.025574691593647003
step: 70, loss: 0.002455455018207431
step: 80, loss: 0.0006641654763370752
step: 90, loss: 0.004092234652489424
step: 100, loss: 0.05843345820903778
step: 110, loss: 0.008941650390625
step: 120, loss: 0.002575373277068138
step: 130, loss: 0.008146314881742
step: 140, loss: 0.047392118722200394
step: 150, loss: 0.03832494840025902
step: 160, loss: 0.02536737360060215
step: 170, loss: 0.13088582456111908
step: 180, loss: 0.01079919096082449
step: 190, loss: 0.02175682783126831
step: 200, loss: 0.03208165988326073
step: 210, loss: 0.07283413410186768
step: 220, loss: 0.01658507250249386
step: 230, loss: 0.027871673926711082
step: 240, loss: 0.0010413771960884333
step: 250, loss: 0.014236735180020332
step: 260, loss: 0.00879299733787775
step: 270, loss: 0.007701556663960218
step: 280, loss: 0.10335098206996918
step: 290, loss: 0.0023200330324470997
step: 300, loss: 0.008407210931181908
step: 310, loss: 0.005493192933499813
step: 320, loss: 0.045306261628866196
step: 330, loss: 0.006705512758344412
step: 340, loss: 0.0019359278958290815
step: 350, loss: 0.0020974576473236084
step: 360, loss: 0.03912178426980972
step: 370, loss: 0.00956891942769289
step: 380, loss: 0.03685806319117546
step: 390, loss: 0.021263204514980316
step: 400, loss: 0.0019905928056687117
step: 410, loss: 0.004189360421150923
step: 420, loss: 0.2125537246465683
step: 430, loss: 0.006516400724649429
step: 440, loss: 0.004996203817427158
step: 450, loss: 0.05769738182425499
step: 460, loss: 0.048455044627189636
step: 470, loss: 0.017009617760777473
step: 480, loss: 0.005313524976372719

```

step: 490, loss: 0.01269361563026905
step: 500, loss: 0.026642754673957825
step: 510, loss: 0.0033827456645667553
step: 520, loss: 0.0984954759478569
step: 530, loss: 0.032085202634334564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9477124183006537, f1=0.9399161620866325,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024613836780190468
step: 10, loss: 0.008888543583452702
step: 20, loss: 0.0008834009058773518
step: 30, loss: 0.0007159108063206077
step: 40, loss: 0.0007518671336583793
step: 50, loss: 0.01045437715947628
step: 60, loss: 0.0009965570643544197
step: 70, loss: 0.009171978570520878
step: 80, loss: 0.009744248352944851
step: 90, loss: 0.0486360602080822
step: 100, loss: 0.002425159327685833
step: 110, loss: 0.0020186398178339005
step: 120, loss: 0.000408641790272668
step: 130, loss: 0.0021698956843465567
step: 140, loss: 0.0007574913324788213
step: 150, loss: 0.010417554527521133
step: 160, loss: 0.04638514667749405
step: 170, loss: 0.002773378510028124
step: 180, loss: 0.0077866483479738235
step: 190, loss: 0.002537826308980584
step: 200, loss: 0.0042170193046331406
step: 210, loss: 0.028916481882333755
step: 220, loss: 0.0011889764573425055
step: 230, loss: 0.11392736434936523
step: 240, loss: 0.003718641819432378
step: 250, loss: 0.001440187799744308
step: 260, loss: 0.008688585832715034
step: 270, loss: 0.011753340251743793
step: 280, loss: 0.04948883503675461
step: 290, loss: 0.007458096835762262
step: 300, loss: 0.00015441107098013163

```

```

step: 310, loss: 0.0431721955537796
step: 320, loss: 0.003918673377484083
step: 330, loss: 0.0030123356264084578
step: 340, loss: 0.006176785100251436
step: 350, loss: 0.01922823116183281
step: 360, loss: 0.03611089289188385
step: 370, loss: 0.0036561116576194763
step: 380, loss: 0.02032882533967495
step: 390, loss: 0.0006469326326623559
step: 400, loss: 0.0007488298579119146
step: 410, loss: 0.007777323015034199
step: 420, loss: 0.0012085551861673594
step: 430, loss: 0.10118141025304794
step: 440, loss: 0.0015196712920442224
step: 450, loss: 0.001815195195376873
step: 460, loss: 0.0007057341863401234
step: 470, loss: 0.0218212828040123
step: 480, loss: 0.048717398196458817
step: 490, loss: 0.0022155302576720715
step: 500, loss: 0.018251821398735046
step: 510, loss: 0.007748677395284176
step: 520, loss: 0.0791904628276825
step: 530, loss: 0.02425110526382923
/usr/local/lib/python3.7/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=0.9421564090698751, f1=0.9373549883990719,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015403380617499352
step: 10, loss: 0.05983789637684822
step: 20, loss: 0.003955284599214792
step: 30, loss: 0.00035320603637956083
step: 40, loss: 0.003438769606873393
step: 50, loss: 0.0002358785131946206
step: 60, loss: 0.02231016755104065
step: 70, loss: 0.0020640722941607237
step: 80, loss: 0.00012094805424567312
step: 90, loss: 0.0008872812031768262
step: 100, loss: 0.02557271346449852
step: 110, loss: 0.00013206420408096164
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 120, loss: 0.017217569053173065
step: 130, loss: 0.00030171076650731266
step: 140, loss: 0.0008375071338377893
step: 150, loss: 0.031198883429169655

```

```

step: 160, loss: 0.004835787694901228
step: 170, loss: 0.05558237433433533
step: 180, loss: 0.0014264322817325592
step: 190, loss: 0.0006914436817169189
step: 200, loss: 0.0007524079410359263
step: 210, loss: 0.006812218576669693
step: 220, loss: 0.000782727962359786
step: 230, loss: 0.0020549409091472626
step: 240, loss: 0.0034773764200508595
step: 250, loss: 0.0006172394496388733
step: 260, loss: 0.0018232985166832805
step: 270, loss: 0.0006850066711194813
step: 280, loss: 0.0031613088212907314
step: 290, loss: 0.1082211434841156
step: 300, loss: 0.0025163660757243633
step: 310, loss: 0.010167674161493778
step: 320, loss: 0.03458520025014877
step: 330, loss: 0.005454624071717262
step: 340, loss: 0.0022259922698140144
step: 350, loss: 0.00038461238727904856
step: 360, loss: 0.007524747401475906
step: 370, loss: 0.055160604417324066
step: 380, loss: 0.00030838287784717977
step: 390, loss: 0.00018759883823804557
step: 400, loss: 0.013909811154007912
step: 410, loss: 0.0019123721867799759
step: 420, loss: 0.00387693103402853
step: 430, loss: 0.007152121514081955
step: 440, loss: 0.017888307571411133
step: 450, loss: 0.004910976625978947
step: 460, loss: 0.0004569849115796387
step: 470, loss: 0.0016232527559623122
step: 480, loss: 0.039910703897476196
step: 490, loss: 0.0006482237367890775
step: 500, loss: 0.009265676140785217
step: 510, loss: 0.016092516481876373
step: 520, loss: 0.0004594168276526034
step: 530, loss: 0.004576984792947769
/usr/local/lib/python3.7/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=0.9398601398601399, f1=0.9355900329102022,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005055416841059923
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```


step: 10, loss: 0.003125089220702648
step: 20, loss: 0.11836227774620056
step: 30, loss: 0.0016770302318036556
step: 40, loss: 0.00216876994818449
step: 50, loss: 0.024211114272475243
step: 60, loss: 0.00013479309563990682
step: 70, loss: 0.0002557132102083415
step: 80, loss: 0.00019745116878766567
step: 90, loss: 0.005089359823614359
step: 100, loss: 0.027321938425302505
step: 110, loss: 0.007326062303036451
step: 120, loss: 0.0009189686388708651
step: 130, loss: 0.003928322345018387
step: 140, loss: 0.0130965830758214
step: 150, loss: 0.0003544864011928439
step: 160, loss: 0.003202717751264572
step: 170, loss: 0.00017592923541087657
step: 180, loss: 0.00046329459291882813
step: 190, loss: 0.0004335502162575722
step: 200, loss: 0.00039979981374926865
step: 210, loss: 0.0010734500829130411
step: 220, loss: 0.002521219663321972
step: 230, loss: 0.0017752374988049269
step: 240, loss: 0.09441922605037689
step: 250, loss: 0.00300008081831038
step: 260, loss: 0.0002764519304037094
step: 270, loss: 0.05748853459954262
step: 280, loss: 0.0053177280351519585
step: 290, loss: 0.00048595500993542373
step: 300, loss: 0.003878458635881543
step: 310, loss: 0.0039365049451589584
step: 320, loss: 0.00024201953783631325
step: 330, loss: 0.00027312347083352506
step: 340, loss: 0.04921768233180046
step: 350, loss: 0.0024652034044265747
step: 360, loss: 0.010858186520636082
step: 370, loss: 0.0028127734549343586
step: 380, loss: 0.0002255178987979889
step: 390, loss: 0.009278016164898872
step: 400, loss: 0.0004719517019111663
step: 410, loss: 0.0010849563404917717
step: 420, loss: 0.002184014767408371
step: 430, loss: 0.007665320299565792
step: 440, loss: 0.0033478830009698868
step: 450, loss: 0.0002464276331011206
step: 460, loss: 0.00015061911835800856
step: 470, loss: 0.010563332587480545
step: 480, loss: 0.02667563036084175

```

step: 490, loss: 0.006535269785672426
step: 500, loss: 0.002525540068745613
step: 510, loss: 0.0018067555502057076
step: 520, loss: 0.0006948797381483018
step: 530, loss: 0.005464083049446344
/usr/local/lib/python3.7/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=0.943778801843318, f1=0.9455216989843028,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021345894783735275
step: 10, loss: 0.0010638275416567922
step: 20, loss: 0.008050945587456226
step: 30, loss: 0.007171256002038717
step: 40, loss: 0.042330302298069
step: 50, loss: 0.002630645642057061
step: 60, loss: 0.006121674086898565
step: 70, loss: 0.0008913237834349275
step: 80, loss: 0.04872102290391922
step: 90, loss: 0.003330082865431905
step: 100, loss: 0.00018099317094311118
step: 110, loss: 0.000982042751275003
step: 120, loss: 0.00019725775928236544
step: 130, loss: 0.03212897852063179
step: 140, loss: 0.00023713284463156015
step: 150, loss: 0.0010888184187933803
step: 160, loss: 0.0003707190044224262
step: 170, loss: 0.000441904179751873
step: 180, loss: 0.0007657140959054232
step: 190, loss: 0.0008899106760509312
step: 200, loss: 0.0009700088994577527
step: 210, loss: 0.0007775938720442355
step: 220, loss: 8.826475823298097e-05
step: 230, loss: 0.03464944660663605
step: 240, loss: 0.0016239839605987072
step: 250, loss: 0.002066037617623806
step: 260, loss: 0.00021492282394319773
step: 270, loss: 0.00013943665544502437
step: 280, loss: 0.0008684640051797032
step: 290, loss: 0.00014650626690126956
step: 300, loss: 0.02146965265274048
step: 310, loss: 0.00014840747462585568
step: 320, loss: 0.0015146504156291485
step: 330, loss: 0.03128195181488991
step: 340, loss: 0.04249422997236252

```

```

step: 350, loss: 0.0010689005721360445
step: 360, loss: 0.0018164144130423665
step: 370, loss: 0.0005698201130144298
step: 380, loss: 0.013423648662865162
step: 390, loss: 0.007238948717713356
step: 400, loss: 0.024282747879624367
step: 410, loss: 0.0010378031292930245
step: 420, loss: 0.0021682048682123423
step: 430, loss: 7.56757945055142e-05
step: 440, loss: 0.01939702033996582
step: 450, loss: 0.0005183693720027804
step: 460, loss: 0.0004182447446510196
step: 470, loss: 0.025638829916715622
step: 480, loss: 0.009808368980884552
step: 490, loss: 0.0008435165509581566
step: 500, loss: 0.016805900260806084
step: 510, loss: 0.0015084930928424
step: 520, loss: 0.02125234343111515
step: 530, loss: 0.0002923100837506354
/usr/local/lib/python3.7/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=0.9389799635701275, f1=0.9276556776556777,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007690052967518568
step: 10, loss: 0.08169148117303848
step: 20, loss: 0.015062396414577961
step: 30, loss: 0.003820492187514901
step: 40, loss: 0.00035063494578935206
step: 50, loss: 0.005749600939452648
step: 60, loss: 0.0009620976634323597
step: 70, loss: 0.00035189263871870935
step: 80, loss: 0.0040920330211520195
step: 90, loss: 0.0006496134446933866
step: 100, loss: 0.00014203709724824876
step: 110, loss: 0.015998654067516327
step: 120, loss: 0.00010685676534194499
step: 130, loss: 0.008477688767015934
step: 140, loss: 0.004923657048493624
step: 150, loss: 0.0002421543322270736
step: 160, loss: 0.001972633646801114
step: 170, loss: 0.001466096262447536
step: 180, loss: 0.0002132585213985294
step: 190, loss: 0.0023379516787827015
step: 200, loss: 0.0010448024841025472

```

```

step: 210, loss: 0.004335293546319008
step: 220, loss: 0.0010803723707795143
step: 230, loss: 0.00019353836250957102
step: 240, loss: 0.0020723638590425253
step: 250, loss: 0.00018023696611635387
step: 260, loss: 0.00012372495257295668
step: 270, loss: 0.03123420849442482
step: 280, loss: 0.008535858243703842
step: 290, loss: 0.0026330731343477964
step: 300, loss: 0.0002455335925333202
step: 310, loss: 0.001390489749610424
step: 320, loss: 0.00823777262121439
step: 330, loss: 0.17045311629772186
step: 340, loss: 0.009179625660181046
step: 350, loss: 0.00018181808991357684
step: 360, loss: 0.0004670696216635406
step: 370, loss: 0.000410695793107152
step: 380, loss: 0.0006754281930625439
step: 390, loss: 0.1256365180015564
step: 400, loss: 8.529525075573474e-05
step: 410, loss: 0.00028261233819648623
step: 420, loss: 0.0020656720735132694
step: 430, loss: 0.04964274540543556
step: 440, loss: 0.00016732177755329758
step: 450, loss: 0.004702033940702677
step: 460, loss: 0.00015447914483956993
step: 470, loss: 0.00012723305553663522
step: 480, loss: 7.452531281160191e-05
step: 490, loss: 0.0018553276313468814
step: 500, loss: 0.002650592476129532
step: 510, loss: 0.0003889107902068645
step: 520, loss: 0.00729097006842494
step: 530, loss: 0.00024200548068620265
/usr/local/lib/python3.7/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=0.9433611884865365, f1=0.9411764705882353,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006573043065145612
step: 10, loss: 4.558268847176805e-05
step: 20, loss: 0.0005167967756278813
step: 30, loss: 0.00012998326565138996
step: 40, loss: 6.78995784255676e-05
step: 50, loss: 3.628231934271753e-05
step: 60, loss: 6.537616718560457e-05

```

step: 70, loss: 7.598072988912463e-05
step: 80, loss: 0.0007954614702612162
step: 90, loss: 7.82223287387751e-05
step: 100, loss: 6.653562013525516e-05
step: 110, loss: 4.994788469048217e-05
step: 120, loss: 9.277651406591758e-05
step: 130, loss: 0.0036835370119661093
step: 140, loss: 0.00011235721467528492
step: 150, loss: 6.045601185178384e-05
step: 160, loss: 0.0018800945254042745
step: 170, loss: 0.008279784582555294
step: 180, loss: 4.1315011912956834e-05
step: 190, loss: 0.008353604935109615
step: 200, loss: 0.00013313234376255423
step: 210, loss: 7.868052489357069e-05
step: 220, loss: 0.00488616107031703
step: 230, loss: 0.0008984058513306081
step: 240, loss: 0.0001776267308741808
step: 250, loss: 0.001871780608780682
step: 260, loss: 0.003246503882110119
step: 270, loss: 2.7495574613567442e-05
step: 280, loss: 6.388730253092945e-05
step: 290, loss: 0.05959200859069824
step: 300, loss: 4.4364020141074434e-05
step: 310, loss: 0.0539490282535553
step: 320, loss: 0.00026606765459291637
step: 330, loss: 0.00033100286964327097
step: 340, loss: 7.861690392019227e-05
step: 350, loss: 0.00013320904690772295
step: 360, loss: 7.483959780074656e-05
step: 370, loss: 0.04238790273666382
step: 380, loss: 7.262352301040664e-05
step: 390, loss: 3.3160926250275224e-05
step: 400, loss: 5.791310468339361e-05
step: 410, loss: 2.7502921511768363e-05
step: 420, loss: 2.2023597921361215e-05
step: 430, loss: 6.60957521176897e-05
step: 440, loss: 0.0006720696110278368
step: 450, loss: 2.859811320377048e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 460, loss: 8.60485015437007e-05
step: 470, loss: 1.8942815586342476e-05
step: 480, loss: 3.159650441375561e-05
step: 490, loss: 0.0018427754985168576
step: 500, loss: 0.0004335816774982959
step: 510, loss: 5.9197162045165896e-05
step: 520, loss: 2.640798265929334e-05
step: 530, loss: 0.0005667585646733642

```

/usr/local/lib/python3.7/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=0.9465861588481189, f1=0.9433085501858736,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.51449520571623e-05
step: 10, loss: 2.5371919036842883e-05
step: 20, loss: 2.3375499949906953e-05
step: 30, loss: 2.7685591703630053e-05
step: 40, loss: 7.897821342339739e-05
step: 50, loss: 0.00010081043001264334
step: 60, loss: 4.599761086865328e-05
step: 70, loss: 5.246422369964421e-05
step: 80, loss: 2.0950697944499552e-05
step: 90, loss: 0.012563930824398994
step: 100, loss: 3.516626748023555e-05
step: 110, loss: 1.790353599062655e-05
step: 120, loss: 0.0013247043825685978
step: 130, loss: 9.808153117774054e-05
step: 140, loss: 0.0008672905387356877
step: 150, loss: 3.12864976876881e-05
step: 160, loss: 2.4820908947731368e-05
step: 170, loss: 3.207010013284162e-05
step: 180, loss: 2.7078618586529046e-05
step: 190, loss: 0.000244640075834468
step: 200, loss: 0.002591117052361369
step: 210, loss: 4.920561332255602e-05
step: 220, loss: 0.00017246014613192528
step: 230, loss: 0.0018569901585578918
step: 240, loss: 2.8791744625777937e-05
step: 250, loss: 3.711605677381158e-05
step: 260, loss: 0.0018963165348395705
step: 270, loss: 0.000112898153020069
step: 280, loss: 2.431448592687957e-05
step: 290, loss: 0.001280934433452785
step: 300, loss: 1.8220118363387883e-05
step: 310, loss: 3.958563684136607e-05
step: 320, loss: 0.00786388386040926
step: 330, loss: 2.4756389393587597e-05
step: 340, loss: 0.00015476264525204897
step: 350, loss: 4.203770004096441e-05
step: 360, loss: 3.0483066439046524e-05
step: 370, loss: 0.0020136404782533646
step: 380, loss: 0.00011348981206538156
step: 390, loss: 5.143459202372469e-05

```

```

step: 400, loss: 0.0010631275363266468
step: 410, loss: 0.0024191911797970533
step: 420, loss: 0.0004695510433521122
step: 430, loss: 2.260080145788379e-05
step: 440, loss: 0.002876382553949952
step: 450, loss: 2.569628122728318e-05
step: 460, loss: 5.948927719146013e-05
step: 470, loss: 7.257470861077309e-05
step: 480, loss: 2.5852950784610584e-05
step: 490, loss: 3.422637382755056e-05
step: 500, loss: 0.003990327939391136
step: 510, loss: 2.6872970920521766e-05
step: 520, loss: 4.3575888412306085e-05
step: 530, loss: 0.0011774409795179963
/usr/local/lib/python3.7/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=0.9444958371877892, f1=0.9410672853828307,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017669436056166887
step: 10, loss: 2.613586184452288e-05
step: 20, loss: 3.4937118471134454e-05
step: 30, loss: 4.864819857175462e-05
step: 40, loss: 2.8206866772961803e-05
step: 50, loss: 0.00015429267659783363
step: 60, loss: 0.012887670658528805
step: 70, loss: 2.2440648535848595e-05
step: 80, loss: 6.170967390062287e-05
step: 90, loss: 3.066462522838265e-05
step: 100, loss: 2.0160841813776642e-05
step: 110, loss: 4.094234827789478e-05
step: 120, loss: 1.8138118321076035e-05
step: 130, loss: 1.6428302842541598e-05
step: 140, loss: 2.6023648388218135e-05
step: 150, loss: 1.802981751097832e-05
step: 160, loss: 2.4265729734906927e-05
step: 170, loss: 2.0458925064303912e-05
step: 180, loss: 4.300364525988698e-05
step: 190, loss: 5.911951302550733e-05
step: 200, loss: 0.004015162121504545
step: 210, loss: 4.42880300397519e-05
step: 220, loss: 0.0012107104994356632
step: 230, loss: 1.428261475666659e-05
step: 240, loss: 0.0001291835942538455
step: 250, loss: 1.7217973436345346e-05

```

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 260, loss: 4.334102413849905e-05
step: 270, loss: 4.673949661082588e-05
step: 280, loss: 6.448041676776484e-05
step: 290, loss: 3.429612479521893e-05
step: 300, loss: 2.1814807041664608e-05
step: 310, loss: 2.2380860173143446e-05
step: 320, loss: 2.271994344482664e-05
step: 330, loss: 6.962721818126738e-05
step: 340, loss: 2.483177013345994e-05
step: 350, loss: 2.03543204406742e-05
step: 360, loss: 1.8868184270104393e-05
step: 370, loss: 0.0010490284767001867
step: 380, loss: 1.7832730009104125e-05
step: 390, loss: 2.1233356164884754e-05
step: 400, loss: 2.1047217160230502e-05
step: 410, loss: 1.9993294699816033e-05
step: 420, loss: 4.621214247890748e-05
step: 430, loss: 2.3014246835373342e-05
step: 440, loss: 1.754213371896185e-05
step: 450, loss: 2.368088098592125e-05
step: 460, loss: 0.00024678371846675873
step: 470, loss: 1.915139910124708e-05
step: 480, loss: 0.00034171933657489717
step: 490, loss: 2.316280733793974e-05
step: 500, loss: 0.00012919970322400331
step: 510, loss: 2.2559250282938592e-05
step: 520, loss: 7.379525050055236e-05
step: 530, loss: 1.930035614350345e-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=0.9391056137012369, f1=0.9313632030505243,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004673769581131637
step: 10, loss: 2.6619132768246345e-05
step: 20, loss: 7.983499381225556e-05
step: 30, loss: 0.007546865381300449
step: 40, loss: 0.00013573335309047252
step: 50, loss: 0.09472114592790604
step: 60, loss: 0.0009889136999845505
step: 70, loss: 7.691401697229594e-05
step: 80, loss: 9.511301323072985e-05
step: 90, loss: 0.0004924812237732112
step: 100, loss: 6.0071666666772217e-05

```



```
step: 110, loss: 3.9813865441828966e-05
step: 120, loss: 8.205714402720332e-05
step: 130, loss: 0.0007740080473013222
step: 140, loss: 7.874474977143109e-05
step: 150, loss: 0.0004005843074992299
step: 160, loss: 2.09542613447411e-05
step: 170, loss: 2.269017932121642e-05
step: 180, loss: 6.300244422163814e-05
step: 190, loss: 9.79854230536148e-05
step: 200, loss: 3.488889706204645e-05
step: 210, loss: 0.00013199204113334417
step: 220, loss: 3.712097532115877e-05
step: 230, loss: 1.6942318325163797e-05
step: 240, loss: 4.5353957830229774e-05
step: 250, loss: 0.0001362785551464185
step: 260, loss: 0.00010434830619487911
step: 270, loss: 0.00014091251068748534
step: 280, loss: 0.0012606899254024029
step: 290, loss: 0.00011554932279977947
step: 300, loss: 0.002940305508673191
step: 310, loss: 3.387115793884732e-05
step: 320, loss: 2.0756526282639243e-05
step: 330, loss: 9.833479271037504e-05
step: 340, loss: 2.3330907424679026e-05
step: 350, loss: 5.2999625040683895e-05
step: 360, loss: 0.0009175242157652974
step: 370, loss: 7.02356337569654e-05
step: 380, loss: 6.315850623650476e-05
step: 390, loss: 0.0001861492928583175
step: 400, loss: 0.008677978068590164
step: 410, loss: 0.00037222495302557945
step: 420, loss: 0.005002320744097233
step: 430, loss: 0.02543305605649948
step: 440, loss: 3.289937740191817e-05
step: 450, loss: 5.679131209035404e-05
step: 460, loss: 0.00014408478455152363
step: 470, loss: 2.2418143998947926e-05
step: 480, loss: 3.876944902003743e-05
step: 490, loss: 6.866417970741168e-05
step: 500, loss: 0.0005238772719167173
step: 510, loss: 0.0009733563638292253
step: 520, loss: 3.96138675569091e-05
step: 530, loss: 0.00015041801088955253
/usr/local/lib/python3.7/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=0.9435370975268316, f1=0.9426573426573427,
best_f1=0.9399161620866325
```

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

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

step: 0, loss: 0.0007901164353825152
step: 10, loss: 2.295460035384167e-05
step: 20, loss: 1.719552710710559e-05
step: 30, loss: 6.999931065365672e-05
step: 40, loss: 2.9432238079607487e-05
step: 50, loss: 4.89221238240134e-05
step: 60, loss: 2.2630310922977515e-05
step: 70, loss: 1.7460077287978493e-05
step: 80, loss: 1.6614543710602447e-05
step: 90, loss: 0.0071747018955647945
step: 100, loss: 0.0011459162924438715
step: 110, loss: 4.11032306146808e-05
step: 120, loss: 2.4135248168022372e-05
step: 130, loss: 1.4122365428193007e-05
step: 140, loss: 9.051237429957837e-05
step: 150, loss: 0.00028299252153374255
step: 160, loss: 2.1922334781265818e-05
step: 170, loss: 2.0168250557617284e-05
step: 180, loss: 0.0032287712674587965
step: 190, loss: 3.2188214390771464e-05
step: 200, loss: 2.3472210159525275e-05
step: 210, loss: 2.0968998796888627e-05
step: 220, loss: 3.7459409213624895e-05
step: 230, loss: 1.3068165571894497e-05
step: 240, loss: 0.00013084683450870216
step: 250, loss: 2.117027725034859e-05
step: 260, loss: 1.6707563190720975e-05
step: 270, loss: 1.5683224773965776e-05
step: 280, loss: 2.5498720788164064e-05
step: 290, loss: 2.109157867380418e-05
step: 300, loss: 8.493608038406819e-05
step: 310, loss: 0.0002540223940741271
step: 320, loss: 0.00414494052529335
step: 330, loss: 0.0014156863326206803
step: 340, loss: 6.182343349792063e-05
step: 350, loss: 0.017368417233228683
step: 360, loss: 2.0719411622849293e-05
step: 370, loss: 3.437503255554475e-05
step: 380, loss: 0.00022583802638109773
step: 390, loss: 0.00010152206232305616
step: 400, loss: 2.787106677715201e-05
step: 410, loss: 0.0006467009661719203
step: 420, loss: 6.45244654151611e-05
step: 430, loss: 0.006811109837144613
step: 440, loss: 3.347950769239105e-05

```

step: 450, loss: 6.631792348343879e-05
step: 460, loss: 0.0018794428324326873
step: 470, loss: 1.7463549738749862e-05
step: 480, loss: 5.83787405048497e-05
step: 490, loss: 2.0861187294940464e-05
step: 500, loss: 2.3214857719722204e-05
step: 510, loss: 8.03159418865107e-05
step: 520, loss: 0.00011594466195674613
step: 530, loss: 1.880486342997756e-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=0.9457720588235294, f1=0.9452369995398067,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045912369387224317
step: 10, loss: 1.4327276403491851e-05
step: 20, loss: 0.0002831645542755723
step: 30, loss: 3.654470492620021e-05
step: 40, loss: 0.00019057821191381663
step: 50, loss: 6.293827027548105e-05
step: 60, loss: 4.031746357213706e-05
step: 70, loss: 5.030916872783564e-05
step: 80, loss: 2.5062639906536788e-05
step: 90, loss: 0.00020107108866795897
step: 100, loss: 0.0001536735362606123
step: 110, loss: 0.002240272704511881
step: 120, loss: 2.1170117179281078e-05
step: 130, loss: 0.021870722994208336
step: 140, loss: 4.394792267703451e-05
step: 150, loss: 1.7005420886562206e-05
step: 160, loss: 0.0006702140672132373
step: 170, loss: 0.0007720626308582723
step: 180, loss: 1.652885475778021e-05
step: 190, loss: 4.737552808364853e-05
step: 200, loss: 8.303460344905034e-05
step: 210, loss: 2.2403282855520956e-05
step: 220, loss: 1.6301630239468068e-05
step: 230, loss: 0.0010846074437722564
step: 240, loss: 2.0950346879544668e-05
step: 250, loss: 0.0015688241692259908
step: 260, loss: 9.922437311615795e-05
step: 270, loss: 1.671131030889228e-05
step: 280, loss: 1.4528423889714759e-05
step: 290, loss: 0.00016036360466387123
step: 300, loss: 2.4876222596503794e-05

```

```

step: 310, loss: 0.00011012928007403389
step: 320, loss: 6.883824244141579e-05
step: 330, loss: 5.07275435666088e-05
step: 340, loss: 2.983755257446319e-05
step: 350, loss: 0.0001062895025825128
step: 360, loss: 0.00038032865268178284
step: 370, loss: 2.4190565454773605e-05
step: 380, loss: 1.8249755157739855e-05
step: 390, loss: 0.017764948308467865
step: 400, loss: 0.00016178675286937505
step: 410, loss: 1.7564194422448054e-05
step: 420, loss: 1.3980714356875978e-05
step: 430, loss: 5.137327025295235e-05
step: 440, loss: 5.075664375908673e-05
step: 450, loss: 4.472918953979388e-05
step: 460, loss: 3.3523363526910543e-05
step: 470, loss: 1.5180350601440296e-05
step: 480, loss: 2.0339393813628703e-05
step: 490, loss: 2.7479749405756593e-05
step: 500, loss: 1.4286196346802171e-05
step: 510, loss: 9.067555947694927e-05
step: 520, loss: 2.8675416615442373e-05
step: 530, loss: 4.4395190343493596e-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=0.9388136384866884, f1=0.9393656716417911,
best_f1=0.9399161620866325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.3809333621757105e-05
step: 10, loss: 3.280344753875397e-05
step: 20, loss: 0.00017695664428174496
step: 30, loss: 2.0126872186665423e-05
step: 40, loss: 0.03795396909117699
step: 50, loss: 6.224629032658413e-05
step: 60, loss: 1.602178053872194e-05
step: 70, loss: 6.198582559591159e-05
step: 80, loss: 1.5683250239817426e-05
step: 90, loss: 1.75384138856316e-05
step: 100, loss: 3.210917202522978e-05
step: 110, loss: 7.322147575905547e-05
step: 120, loss: 2.959382618428208e-05
step: 130, loss: 0.06480041146278381
step: 140, loss: 1.4390525393537246e-05
step: 150, loss: 2.7528658392839134e-05
step: 160, loss: 2.5422907128813677e-05

```

```

step: 170, loss: 1.8335433196625672e-05
step: 180, loss: 1.0941090295091271e-05
step: 190, loss: 2.0507348381215706e-05
step: 200, loss: 4.561386231216602e-05
step: 210, loss: 4.12810331908986e-05
step: 220, loss: 1.5820927728782408e-05
step: 230, loss: 0.00031953331199474633
step: 240, loss: 1.886802783701569e-05
step: 250, loss: 1.9445140424068086e-05
step: 260, loss: 3.5145028959959745e-05
step: 270, loss: 1.9188388250768185e-05
step: 280, loss: 2.658116318343673e-05
step: 290, loss: 1.1447717952250969e-05
step: 300, loss: 2.449175553920213e-05
step: 310, loss: 1.8506882042856887e-05
step: 320, loss: 1.3865325854567345e-05
step: 330, loss: 0.00012860707647632807
step: 340, loss: 1.1820234249171335e-05
step: 350, loss: 1.2039981811540201e-05
step: 360, loss: 0.001206144574098289
step: 370, loss: 2.9307122531463392e-05
step: 380, loss: 1.798540324671194e-05
step: 390, loss: 4.0007682400755584e-05
step: 400, loss: 1.7367035979987122e-05
step: 410, loss: 1.440924188500503e-05
step: 420, loss: 1.6972142475424334e-05
step: 430, loss: 0.00024362037947867066
step: 440, loss: 3.037416172446683e-05
step: 450, loss: 1.2185266314190812e-05
step: 460, loss: 0.00010243855649605393
step: 470, loss: 0.0010772290406748652
step: 480, loss: 2.200638482463546e-05
step: 490, loss: 1.356734810542548e-05
step: 500, loss: 2.569583557487931e-05
step: 510, loss: 1.270684788323706e-05
step: 520, loss: 1.431984765076777e-05
step: 530, loss: 3.0094775866018608e-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=0.9402023919043238, f1=0.9419354838709678,
best_f1=0.9399161620866325
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled          : True

```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 238.17it/s]
load_f1 = 0.9485981308411215
real_f1 = 0.9472701819878675
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 [00:17, 244.52it/s]

```

3 DK

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                               | 827.9 MB 1.2 MB/s
Requirement already satisfied: spacy>=2.2.2 in

```

```

/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: 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: 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: 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: 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: 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: 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: 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: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: 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: 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: 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: 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=b807871df0b84c8071c342cdfff04ea3c4030a2be23e5a3c020f52d72ae836a2
  Stored in directory: /tmp/pip-ephem-wheel-cache-u3hplcqz/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

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

```



```

--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5505433082580566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.3146067415730337, f1=0.3218390804597701,
best_f1=0.3218390804597701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5107592940330505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.39999999999999997, f1=0.3728813559322034,
best_f1=0.3728813559322034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5641175508499146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4864864864864865, f1=0.43902439024390244,
best_f1=0.43902439024390244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26819559931755066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.6363636363636364, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23554660379886627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7096774193548386, 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: 0.2009439468383789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.782608695652174, 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: 0.037566836923360825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7567567567567568,
best_f1=0.7567567567567568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010561340488493443

```

```

/usr/local/lib/python3.7/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=0.6875000000000001,
best_f1=0.7567567567567568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00447699474170804
/usr/local/lib/python3.7/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=0.7222222222222223,
best_f1=0.7567567567567568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031144116073846817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8750000000000001, f1=0.742857142857143,
best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021032593213021755
/usr/local/lib/python3.7/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=0.7647058823529412,
best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035403247456997633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9032258064516129, f1=0.742857142857143,

```

```

best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003160628955811262
/usr/local/lib/python3.7/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=0.9032258064516129, f1=0.742857142857143,
best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004635571036487818
/usr/local/lib/python3.7/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=0.9032258064516129, f1=0.742857142857143,
best_f1=0.742857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003502734936773777
/usr/local/lib/python3.7/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=0.9032258064516129, f1=0.742857142857143,
best_f1=0.742857142857143
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, 141783.68it/s]
/usr/local/lib/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 = 0.8235294117647058
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 [00:16, 261.41it/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: 0.5912236571311951
```

```
step: 10, loss: 0.613173246383667
```

```
step: 20, loss: 0.30456292629241943
```

```
step: 30, loss: 0.13036657869815826
```

```
step: 40, loss: 0.11315225809812546
```

```

step: 50, loss: 0.015067634172737598
step: 60, loss: 0.08850248157978058
step: 70, loss: 0.1149950698018074
step: 80, loss: 0.10840770602226257
step: 90, loss: 0.08346875756978989
step: 100, loss: 0.004410212859511375
step: 110, loss: 0.1209302693605423
step: 120, loss: 0.004080553073436022
step: 130, loss: 0.009059985168278217
step: 140, loss: 0.0016727561596781015
step: 150, loss: 0.009324206970632076
step: 160, loss: 0.009406063705682755
step: 170, loss: 0.07526449859142303
step: 180, loss: 0.015242502093315125
step: 190, loss: 0.0022401027381420135
step: 200, loss: 0.03187086433172226
step: 210, loss: 0.00333452969789505
step: 220, loss: 0.06455665081739426
step: 230, loss: 0.013138133101165295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932432432432432, f1=0.987598647125141,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019729549530893564
step: 10, loss: 0.0028772568330168724
step: 20, loss: 0.09193333983421326
step: 30, loss: 0.1840655505657196
step: 40, loss: 0.08641267567873001
step: 50, loss: 0.006487919017672539
step: 60, loss: 0.0058038183487951756
step: 70, loss: 0.12218271940946579
step: 80, loss: 0.001207594876177609
step: 90, loss: 0.07038714736700058
step: 100, loss: 0.012712918221950531
step: 110, loss: 0.14255709946155548
step: 120, loss: 0.0029771931003779173
step: 130, loss: 0.014010107144713402
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.010380508378148079
step: 150, loss: 0.0031351863872259855

```



```

step: 160, loss: 0.017583023756742477
step: 170, loss: 0.0011144172167405486
step: 180, loss: 0.003107928903773427
step: 190, loss: 0.01384351309388876
step: 200, loss: 0.002630555536597967
step: 210, loss: 0.0005778797785751522
step: 220, loss: 0.1670699119567871
step: 230, loss: 0.006108736619353294
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9876543209876544,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017101893899962306
step: 10, loss: 0.01818212680518627
step: 20, loss: 0.0008873993647284806
step: 30, loss: 0.0030302382074296474
step: 40, loss: 0.16453196108341217
step: 50, loss: 0.0014561191201210022
step: 60, loss: 0.0030460187699645758
step: 70, loss: 0.007621313910931349
step: 80, loss: 0.01165572740137577
step: 90, loss: 0.09677746891975403
step: 100, loss: 0.001412810175679624
step: 110, loss: 0.0007424347568303347
step: 120, loss: 0.014174762181937695
step: 130, loss: 0.002598279155790806
step: 140, loss: 0.0023226507473737
step: 150, loss: 0.17613892257213593
step: 160, loss: 0.005872709676623344
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.009303612634539604
step: 180, loss: 0.00643333001062274
step: 190, loss: 0.015525607392191887
step: 200, loss: 0.0013971627922728658
step: 210, loss: 0.001922852941788733
step: 220, loss: 0.0010826608631759882
step: 230, loss: 0.004673877265304327
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9865470852017937,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016463998472318053
step: 10, loss: 0.0007781402673572302
step: 20, loss: 0.0011511343764141202
step: 30, loss: 0.0026519543025642633
step: 40, loss: 0.0029359215404838324
step: 50, loss: 0.0005984032759442925
step: 60, loss: 0.0019474721048027277
step: 70, loss: 0.00041076180059462786
step: 80, loss: 0.0011996011016890407
step: 90, loss: 0.0003813990333583206
step: 100, loss: 0.0001909967977553606
step: 110, loss: 0.00025764902238734066
step: 120, loss: 0.00248494790866971
step: 130, loss: 0.004803017247468233
step: 140, loss: 0.0002868474111892283
step: 150, loss: 0.10509353876113892
step: 160, loss: 0.013225874863564968
step: 170, loss: 0.00399406673386693
step: 180, loss: 0.0005847711581736803
step: 190, loss: 0.0028483967762440443
step: 200, loss: 0.0025611722376197577
step: 210, loss: 0.06299956142902374
step: 220, loss: 0.0007206364534795284
step: 230, loss: 0.007565743289887905
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9807909604519773,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010917537845671177
step: 10, loss: 0.0017575782258063555
step: 20, loss: 0.007043652236461639
step: 30, loss: 0.0003917786234524101
step: 40, loss: 0.0004285497998353094
step: 50, loss: 0.000744843331631273
step: 60, loss: 0.0013652791967615485
step: 70, loss: 0.0005708910175599158
step: 80, loss: 0.003000895492732525
step: 90, loss: 0.0005033023771829903
step: 100, loss: 0.005467379465699196
step: 110, loss: 0.0016464312793686986
step: 120, loss: 0.0002515321539249271
step: 130, loss: 0.02526821196079254
step: 140, loss: 0.0007628793246112764

```

```

step: 150, loss: 0.0006527407094836235
step: 160, loss: 0.0007844262290745974
step: 170, loss: 0.010088901966810226
step: 180, loss: 0.021994423121213913
step: 190, loss: 0.015920491889119148
step: 200, loss: 0.0028540620114654303
step: 210, loss: 0.0002701957419048995
step: 220, loss: 0.0003402513393666595
step: 230, loss: 0.0001332952087977901
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9853107344632768,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025089032715186477
step: 10, loss: 0.00021611040574498475
step: 20, loss: 0.0013899814803153276
step: 30, loss: 0.0001642133720451966
step: 40, loss: 0.0006438683485612273
step: 50, loss: 0.00047131121391430497
step: 60, loss: 0.00019868358504027128
step: 70, loss: 0.000495846092235297
step: 80, loss: 0.0025976530741900206
step: 90, loss: 0.00531791104003787
step: 100, loss: 0.03300395607948303
step: 110, loss: 0.0005585573380813003
step: 120, loss: 0.0001974668266484514
step: 130, loss: 0.0002662568585947156
step: 140, loss: 0.000928554916754365
step: 150, loss: 0.000709672924131155
step: 160, loss: 0.00033977848943322897
step: 170, loss: 0.00013082139776088297
step: 180, loss: 0.026050394400954247
step: 190, loss: 0.01797409914433956
step: 200, loss: 9.832880459725857e-05
step: 210, loss: 0.00012599158799275756
step: 220, loss: 0.00024004986335057765
step: 230, loss: 0.06074068322777748
/usr/local/lib/python3.7/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=0.9888392857142857, f1=0.9833147942157954,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.295330578926951e-05
step: 10, loss: 9.747353033162653e-05
step: 20, loss: 0.0003150480333715677
step: 30, loss: 0.0021330092567950487
step: 40, loss: 0.00016454735305160284
step: 50, loss: 0.00024871632922440767
step: 60, loss: 0.0004997088108211756
step: 70, loss: 0.015541023574769497
step: 80, loss: 0.0004044235101900995
step: 90, loss: 6.363772990880534e-05
step: 100, loss: 0.00011655769776552916
step: 110, loss: 8.15505045466125e-05
step: 120, loss: 6.544645293615758e-05
step: 130, loss: 6.0105881857452914e-05
step: 140, loss: 7.668143371120095e-05
step: 150, loss: 0.013339764438569546
step: 160, loss: 0.041934121400117874
step: 170, loss: 0.0010208942694589496
step: 180, loss: 0.0009571127011440694
step: 190, loss: 9.935528214555234e-05
step: 200, loss: 0.027761029079556465
step: 210, loss: 7.700347487116233e-05
step: 220, loss: 0.00021312886383384466
step: 230, loss: 0.00014568993356078863
/usr/local/lib/python3.7/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=0.987709497206704, f1=0.9843749999999999,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.677259418414906e-05
step: 10, loss: 0.012325519695878029
step: 20, loss: 8.598691783845425e-05
step: 30, loss: 0.00037150399293750525
step: 40, loss: 0.0002988656051456928
step: 50, loss: 9.18216974241659e-05
step: 60, loss: 6.166665116325021e-05
step: 70, loss: 9.345063881482929e-05
step: 80, loss: 0.10027477145195007
step: 90, loss: 6.628039409406483e-05
step: 100, loss: 0.00041587871965020895
step: 110, loss: 0.05141633003950119
step: 120, loss: 0.00023868869175203145
step: 130, loss: 0.0006026229821145535
step: 140, loss: 0.00010281162394676358

```

```

step: 150, loss: 0.00019958797201979905
step: 160, loss: 0.0006122904014773667
step: 170, loss: 0.000119932577945292
step: 180, loss: 0.00038917266647331417
step: 190, loss: 0.00010308453056495637
step: 200, loss: 0.00017138360999524593
step: 210, loss: 0.0004793612170033157
step: 220, loss: 0.0012086055940017104
step: 230, loss: 9.625190432416275e-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=0.9909502262443439, f1=0.9831649831649831,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.299958502291702e-05
step: 10, loss: 0.00030645151855424047
step: 20, loss: 0.0008329228730872273
step: 30, loss: 0.0005907490849494934
step: 40, loss: 0.02642367221415043
step: 50, loss: 3.9795228076400235e-05
step: 60, loss: 0.00016424294153694063
step: 70, loss: 0.00010641155677149072
step: 80, loss: 0.00023296572908293456
step: 90, loss: 0.0009939703159034252
step: 100, loss: 0.000757106055971235
step: 110, loss: 0.0006149333203211427
step: 120, loss: 5.2273553592385724e-05
step: 130, loss: 0.00034785782918334007
step: 140, loss: 4.22127268393524e-05
step: 150, loss: 0.0013790099183097482
step: 160, loss: 6.0108468460384756e-05
step: 170, loss: 0.00022803671890869737
step: 180, loss: 0.0014529535546898842
step: 190, loss: 0.00022210771567188203
step: 200, loss: 3.846602339763194e-05
step: 210, loss: 0.0008969073533080518
step: 220, loss: 4.1256000258727e-05
step: 230, loss: 9.608319669496268e-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=0.9886877828054299, f1=0.9864253393665158,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.627736340509728e-05
step: 10, loss: 0.001184658845886588
step: 20, loss: 0.003122099442407489
step: 30, loss: 5.549526758841239e-05
step: 40, loss: 4.782197720487602e-05
step: 50, loss: 4.2390885937493294e-05
step: 60, loss: 0.00044384400825947523
step: 70, loss: 7.379666931228712e-05
step: 80, loss: 5.4427164286607876e-05
step: 90, loss: 0.00014900909445714206
step: 100, loss: 5.725334631279111e-05
step: 110, loss: 9.092612162930891e-05
step: 120, loss: 0.0007230945629999042
step: 130, loss: 7.867210661061108e-05
step: 140, loss: 0.039532847702503204
step: 150, loss: 0.014736157841980457
step: 160, loss: 2.5286759409937076e-05
step: 170, loss: 0.0005395872285589576
step: 180, loss: 0.0018204371444880962
step: 190, loss: 0.0035601570270955563
step: 200, loss: 3.499764352454804e-05
step: 210, loss: 3.817404285655357e-05
step: 220, loss: 5.341934229363687e-05
step: 230, loss: 0.0015736828790977597
/usr/local/lib/python3.7/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=0.9897843359818388, f1=0.9863945578231292,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.264749102527276e-05
step: 10, loss: 0.0003292395849712193
step: 20, loss: 5.2516741561703384e-05
step: 30, loss: 0.021792922168970108
step: 40, loss: 0.0015216651372611523
step: 50, loss: 3.543719503795728e-05
step: 60, loss: 6.392454815795645e-05
step: 70, loss: 0.0004461119242478162
step: 80, loss: 2.5588413336663507e-05
step: 90, loss: 2.4750279408181086e-05
step: 100, loss: 2.9969054594403133e-05
step: 110, loss: 0.01779218763113022
step: 120, loss: 3.2192489015869796e-05
step: 130, loss: 2.3800226699677296e-05
step: 140, loss: 3.988970638602041e-05

```

```

step: 150, loss: 0.024000976234674454
step: 160, loss: 2.9019365683780052e-05
step: 170, loss: 0.023822268471121788
step: 180, loss: 3.384681258467026e-05
step: 190, loss: 2.5882618501782417e-05
step: 200, loss: 3.679515066323802e-05
step: 210, loss: 2.646389475557953e-05
step: 220, loss: 3.448376810410991e-05
step: 230, loss: 0.00014546459715347737
/usr/local/lib/python3.7/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=0.9897843359818388, f1=0.9852440408626559,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.084446143475361e-05
step: 10, loss: 3.291120447102003e-05
step: 20, loss: 2.8929513064213097e-05
step: 30, loss: 3.16230725729838e-05
step: 40, loss: 7.01382668921724e-05
step: 50, loss: 4.159028685535304e-05
step: 60, loss: 0.0006148448446765542
step: 70, loss: 3.10195900965482e-05
step: 80, loss: 5.9524372773012146e-05
step: 90, loss: 2.650474380061496e-05
step: 100, loss: 2.130452594428789e-05
step: 110, loss: 2.981265788548626e-05
step: 120, loss: 3.469971125014126e-05
step: 130, loss: 3.026342164957896e-05
step: 140, loss: 0.02595100924372673
step: 150, loss: 3.742236731341109e-05
step: 160, loss: 3.08968847093638e-05
step: 170, loss: 8.25349343358539e-05
step: 180, loss: 0.08316341787576675
step: 190, loss: 0.0007365595083683729
step: 200, loss: 0.00021999416640028358
step: 210, loss: 6.904733891133219e-05
step: 220, loss: 3.9873346395324916e-05
step: 230, loss: 0.03366339951753616
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9842342342342343,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012079127918696031
step: 10, loss: 0.0005366927944123745
step: 20, loss: 0.00042806490091606975
step: 30, loss: 0.00747909490019083
step: 40, loss: 8.798750059213489e-05
step: 50, loss: 0.0018305268604308367
step: 60, loss: 9.980913455365226e-05
step: 70, loss: 2.4485734684276395e-05
step: 80, loss: 2.983106423926074e-05
step: 90, loss: 0.007402780000120401
step: 100, loss: 2.0794253941858187e-05
step: 110, loss: 0.05620589479804039
step: 120, loss: 0.024302903562784195
step: 130, loss: 0.00017460626258980483
step: 140, loss: 3.0185259674908593e-05
step: 150, loss: 3.484405897324905e-05
step: 160, loss: 0.0008917247760109603
step: 170, loss: 3.904462209902704e-05
step: 180, loss: 9.713986946735531e-05
step: 190, loss: 2.3170832719188184e-05
step: 200, loss: 0.00011212209210498258
step: 210, loss: 1.7218053471879102e-05
step: 220, loss: 2.8192058380227536e-05
step: 230, loss: 2.289518670295365e-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=0.9898305084745763, f1=0.9875706214689265,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6672396415960975e-05
step: 10, loss: 0.013899270445108414
step: 20, loss: 4.014841761090793e-05
step: 30, loss: 4.095615440746769e-05
step: 40, loss: 0.012041827663779259
step: 50, loss: 3.5984834539704025e-05
step: 60, loss: 6.97954383213073e-05
step: 70, loss: 3.49676665791776e-05
step: 80, loss: 3.392072540009394e-05
step: 90, loss: 5.9157613577554e-05
step: 100, loss: 2.8229163945070468e-05
step: 110, loss: 0.000264840287854895
step: 120, loss: 1.662574322836008e-05
step: 130, loss: 2.7626116207102314e-05
step: 140, loss: 7.63141069910489e-05

```



```

step: 150, loss: 4.2038944229716435e-05
step: 160, loss: 0.0028619421645998955
step: 170, loss: 2.1855757950106636e-05
step: 180, loss: 3.285564525867812e-05
step: 190, loss: 9.186744864564389e-05
step: 200, loss: 2.204953671025578e-05
step: 210, loss: 3.936719804187305e-05
step: 220, loss: 0.00014768676192034036
step: 230, loss: 3.0829494789941236e-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=0.9898305084745763, f1=0.9875706214689265,
best_f1=0.987598647125141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6485928174224682e-05
step: 10, loss: 1.5906784028629772e-05
step: 20, loss: 2.818478060362395e-05
step: 30, loss: 2.3967977540451102e-05
step: 40, loss: 4.964138497598469e-05
step: 50, loss: 0.05188938230276108
step: 60, loss: 0.0001014384106383659
step: 70, loss: 0.0005642937612719834
step: 80, loss: 2.3748076273477636e-05
step: 90, loss: 4.0044869820121676e-05
step: 100, loss: 2.8829175789724104e-05
step: 110, loss: 2.5900988475768827e-05
step: 120, loss: 8.50071373861283e-05
step: 130, loss: 0.021541429683566093
step: 140, loss: 1.9590956071624532e-05
step: 150, loss: 5.899393363506533e-05
step: 160, loss: 2.1043677406851202e-05
step: 170, loss: 2.334569398954045e-05
step: 180, loss: 6.476329144788906e-05
step: 190, loss: 6.134119757916778e-05
step: 200, loss: 0.0001101860252674669
step: 210, loss: 2.4046303224167787e-05
step: 220, loss: 2.9052338504698128e-05
step: 230, loss: 2.385251718806103e-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=0.9898305084745763, f1=0.9853438556933484,
best_f1=0.987598647125141
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32

```

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

2473it [00:13, 183.46it/s]

load_f1 = 0.9909706546275394

real_f1 = 0.9887387387387387

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

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

4398it [00:19, 227.56it/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: 0.6350420117378235
step: 10, loss: 0.5049020648002625
step: 20, loss: 0.46452629566192627
step: 30, loss: 0.10125920921564102
step: 40, loss: 0.14172084629535675
step: 50, loss: 0.15689165890216827
step: 60, loss: 0.10903426259756088
step: 70, loss: 0.09903188794851303
step: 80, loss: 0.0835980698466301
step: 90, loss: 0.1248892992734909
step: 100, loss: 0.007366569247096777
step: 110, loss: 0.06017322093248367
step: 120, loss: 0.017051592469215393
step: 130, loss: 0.08691544085741043
step: 140, loss: 0.04463982582092285
step: 150, loss: 0.056360047310590744
step: 160, loss: 0.027388162910938263
step: 170, loss: 0.1821836680173874
step: 180, loss: 0.046608321368694305
step: 190, loss: 0.007845448330044746
step: 200, loss: 0.13117225468158722
step: 210, loss: 0.08506263792514801
step: 220, loss: 0.22120998799800873
step: 230, loss: 0.14998750388622284
step: 240, loss: 0.05285867676138878
step: 250, loss: 0.02568797767162323
step: 260, loss: 0.1153033971786499
step: 270, loss: 0.008436375297605991
step: 280, loss: 0.03697231039404869
step: 290, loss: 0.09269073605537415
step: 300, loss: 0.05288730561733246
step: 310, loss: 0.24970397353172302
step: 320, loss: 0.13029523193836212
step: 330, loss: 0.009609071537852287
step: 340, loss: 0.018440909683704376

```

```

step: 350, loss: 0.06173449009656906
step: 360, loss: 0.025632379576563835
step: 370, loss: 0.04061676561832428
step: 380, loss: 0.021859364584088326
step: 390, loss: 0.2511858642101288
step: 400, loss: 0.16955412924289703
step: 410, loss: 0.057005371898412704
step: 420, loss: 0.015135346911847591
step: 430, loss: 0.16911004483699799
step: 440, loss: 0.008570468053221703
step: 450, loss: 0.006856610998511314
step: 460, loss: 0.006270139943808317
step: 470, loss: 0.17143189907073975
step: 480, loss: 0.07643870264291763
step: 490, loss: 0.08097747713327408
step: 500, loss: 0.06306351721286774
step: 510, loss: 0.0799063891172409
step: 520, loss: 0.02497197315096855
step: 530, loss: 0.007129194214940071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9400749063670412, f1=0.9376163873370577,
best_f1=0.9376163873370577
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2304520606994629
step: 10, loss: 0.04230069741606712
step: 20, loss: 0.024911954998970032
step: 30, loss: 0.017779307439923286
step: 40, loss: 0.18035510182380676
step: 50, loss: 0.18455441296100616
step: 60, loss: 0.009698187932372093
step: 70, loss: 0.04807819426059723
step: 80, loss: 0.06325575709342957
step: 90, loss: 0.014701798558235168
step: 100, loss: 0.005770438350737095
step: 110, loss: 0.0653035119175911
step: 120, loss: 0.029037579894065857
step: 130, loss: 0.09475568681955338
step: 140, loss: 0.014496886171400547
step: 150, loss: 0.061764054000377655
step: 160, loss: 0.007160375826060772

```

```

step: 170, loss: 0.011929324828088284
step: 180, loss: 0.07488108426332474
step: 190, loss: 0.051275115460157394
step: 200, loss: 0.003162249457091093
step: 210, loss: 0.021791014820337296
step: 220, loss: 0.04834854230284691
step: 230, loss: 0.005476110614836216
step: 240, loss: 0.08189613372087479
step: 250, loss: 0.051973339170217514
step: 260, loss: 0.00520080141723156
step: 270, loss: 0.09756011515855789
step: 280, loss: 0.04349921643733978
step: 290, loss: 0.03712889552116394
step: 300, loss: 0.16323134303092957
step: 310, loss: 0.010487556457519531
step: 320, loss: 0.069209985435009
step: 330, loss: 0.01903090812265873
step: 340, loss: 0.0067299711517989635
step: 350, loss: 0.0014702569460496306
step: 360, loss: 0.05519111827015877
step: 370, loss: 0.10310329496860504
step: 380, loss: 0.034561559557914734
step: 390, loss: 0.12216018885374069
step: 400, loss: 0.06742209196090698
step: 410, loss: 0.023544223979115486
step: 420, loss: 0.12178279459476471
step: 430, loss: 0.0077074128203094006
step: 440, loss: 0.16570398211479187
step: 450, loss: 0.007328768260776997
step: 460, loss: 0.10412410646677017
step: 470, loss: 0.2615988850593567
step: 480, loss: 0.3557838797569275
step: 490, loss: 0.017518004402518272
step: 500, loss: 0.14270329475402832
step: 510, loss: 0.012167258188128471
step: 520, loss: 0.04108347371220589
step: 530, loss: 0.02489355579018593
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9438305709023942, f1=0.942634235888022,
best_f1=0.942634235888022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

step: 0, loss: 0.06255937367677689
step: 10, loss: 0.05582696944475174
step: 20, loss: 0.04832927882671356
step: 30, loss: 0.14986862242221832
step: 40, loss: 0.04379558563232422
step: 50, loss: 0.038985252380371094
step: 60, loss: 0.007482155691832304
step: 70, loss: 0.0035661261063069105
step: 80, loss: 0.00090830959379673
step: 90, loss: 0.008147219195961952
step: 100, loss: 0.10530491918325424
step: 110, loss: 0.041972413659095764
step: 120, loss: 0.0030872791539877653
step: 130, loss: 0.0016947611002251506
step: 140, loss: 0.06678985059261322
step: 150, loss: 0.0400523915886879
step: 160, loss: 0.0024607027880847454
step: 170, loss: 0.050212081521749496
step: 180, loss: 0.0022833317052572966
step: 190, loss: 0.006690829060971737
step: 200, loss: 0.0110322255641222
step: 210, loss: 0.10491257905960083
step: 220, loss: 0.0398193784058094
step: 230, loss: 0.08480258285999298
step: 240, loss: 0.010538299567997456
step: 250, loss: 0.01607830822467804
step: 260, loss: 0.01814902201294899
step: 270, loss: 0.0025631634052842855
step: 280, loss: 0.05819517374038696
step: 290, loss: 0.0017809470882639289
step: 300, loss: 0.07956673204898834
step: 310, loss: 0.016780199483036995
step: 320, loss: 0.011824268847703934
step: 330, loss: 0.000613226555287838
step: 340, loss: 0.009801295585930347
step: 350, loss: 0.0008638878935016692
step: 360, loss: 0.029737688601017
step: 370, loss: 0.026474976912140846
step: 380, loss: 0.02035975083708763
step: 390, loss: 0.004253219813108444
step: 400, loss: 0.0892849937081337
step: 410, loss: 0.004057370591908693
step: 420, loss: 0.08993767201900482
step: 430, loss: 0.00227172183804214
step: 440, loss: 0.002486178418621421
step: 450, loss: 0.057112134993076324

```

step: 460, loss: 0.24144814908504486
step: 470, loss: 0.03938394784927368
step: 480, loss: 0.0031058567110449076
step: 490, loss: 0.011798111721873283
step: 500, loss: 0.011466310359537601
step: 510, loss: 0.009260444901883602
step: 520, loss: 0.05141238868236542
step: 530, loss: 0.06967710703611374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9483870967741935, f1=0.9478499542543459,
best_f1=0.9478499542543459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014572839252650738
step: 10, loss: 0.005425744690001011
step: 20, loss: 0.00272850482724607
step: 30, loss: 0.000556253595277667
step: 40, loss: 0.0026805901434272528
step: 50, loss: 0.0015205774689093232
step: 60, loss: 0.0009804086294025183
step: 70, loss: 0.0010012215934693813
step: 80, loss: 0.00021905846369918436
step: 90, loss: 0.027802415192127228
step: 100, loss: 0.0029459381476044655
step: 110, loss: 0.013860419392585754
step: 120, loss: 0.00021782460680697113
step: 130, loss: 0.0004756225971505046
step: 140, loss: 0.0014205550542101264
step: 150, loss: 0.10569918155670166
step: 160, loss: 0.023343201726675034
step: 170, loss: 0.0036125255282968283
step: 180, loss: 0.023640727624297142
step: 190, loss: 0.020025968551635742
step: 200, loss: 0.01823408529162407
step: 210, loss: 0.07943885773420334
step: 220, loss: 0.018966058269143105
step: 230, loss: 0.03028620406985283
step: 240, loss: 0.0016947604017332196
step: 250, loss: 0.045853059738874435
step: 260, loss: 0.0024232210125774145
step: 270, loss: 0.016483677551150322

```



```

step: 280, loss: 0.012056311592459679
step: 290, loss: 0.03107813559472561
step: 300, loss: 9.601706551620737e-05
step: 310, loss: 0.005619543604552746
step: 320, loss: 0.013821030035614967
step: 330, loss: 0.06574956327676773
step: 340, loss: 0.0012723514810204506
step: 350, loss: 0.002897347789257765
step: 360, loss: 0.0898388996720314
step: 370, loss: 0.001358235371299088
step: 380, loss: 0.04796341434121132
step: 390, loss: 0.0014555498491972685
step: 400, loss: 0.0019271252676844597
step: 410, loss: 0.0018894968088716269
step: 420, loss: 0.003237369703128934
step: 430, loss: 0.244017094373703
step: 440, loss: 0.00039349833969026804
step: 450, loss: 0.034653231501579285
step: 460, loss: 0.003563609439879656
step: 470, loss: 0.03244340419769287
step: 480, loss: 0.04141226038336754
step: 490, loss: 0.0014893290353938937
step: 500, loss: 0.00596939492970705
step: 510, loss: 0.006179647520184517
step: 520, loss: 0.03335297480225563
step: 530, loss: 0.020626094192266464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9492455418381345, f1=0.945387792565397,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009518886916339397
step: 10, loss: 0.040077224373817444
step: 20, loss: 0.0005832272581756115
step: 30, loss: 0.002214025938883424
step: 40, loss: 0.0008116030949167907
step: 50, loss: 0.003631901228800416
step: 60, loss: 0.17844147980213165
step: 70, loss: 0.0005070064798928797
step: 80, loss: 0.0005395627813413739
step: 90, loss: 0.0017656831769272685

```

```

step: 100, loss: 0.022235434502363205
step: 110, loss: 0.00047804080531932414
step: 120, loss: 0.0022476189769804478
step: 130, loss: 0.0008641664171591401
step: 140, loss: 0.0005985407624393702
step: 150, loss: 0.004518116824328899
step: 160, loss: 0.00030351089662872255
step: 170, loss: 0.014995378442108631
step: 180, loss: 0.0005146947223693132
step: 190, loss: 0.00018668542907107621
step: 200, loss: 0.0011472422629594803
step: 210, loss: 0.0004117793869227171
step: 220, loss: 0.0013099797070026398
step: 230, loss: 0.0028275675140321255
step: 240, loss: 0.0021397583186626434
step: 250, loss: 0.0005875956267118454
step: 260, loss: 0.014974728226661682
step: 270, loss: 0.0003593654546421021
step: 280, loss: 0.0005191628588363528
step: 290, loss: 0.12586922943592072
step: 300, loss: 0.0017680219607427716
step: 310, loss: 0.004601478576660156
step: 320, loss: 0.06623342633247375
step: 330, loss: 0.016548793762922287
step: 340, loss: 0.012950398959219456
step: 350, loss: 0.0015023527666926384
step: 360, loss: 0.002516228938475251
step: 370, loss: 0.002242650371044874
step: 380, loss: 0.0006989616085775197
step: 390, loss: 0.00015899924619589
step: 400, loss: 0.0012948832008987665
step: 410, loss: 0.0003846421022899449
step: 420, loss: 0.002551198936998844
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 430, loss: 0.11563661694526672
step: 440, loss: 0.008774259127676487
step: 450, loss: 0.17643645405769348
step: 460, loss: 0.008403895422816277
step: 470, loss: 0.014629934914410114
step: 480, loss: 0.0015918738208711147
step: 490, loss: 0.00026200246065855026
step: 500, loss: 0.0025328241754323244
step: 510, loss: 0.0023450569715350866
step: 520, loss: 0.0003787523601204157
step: 530, loss: 0.0005723715876229107
/usr/local/lib/python3.7/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=0.9478584729981377, f1=0.9442896935933148,  
best_f1=0.945387792565397  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0002228762023150921  
step: 10, loss: 0.00013281528663355857  
step: 20, loss: 0.14056390523910522  
step: 30, loss: 0.001854976755566895  
step: 40, loss: 0.0009628336410969496  
step: 50, loss: 0.007042016368359327  
step: 60, loss: 0.0009540215833112597  
step: 70, loss: 0.003373926505446434  
step: 80, loss: 0.0036872997879981995  
step: 90, loss: 0.0038173850625753403  
step: 100, loss: 0.002194071887061  
step: 110, loss: 0.028213901445269585  
step: 120, loss: 0.0003105198557022959  
step: 130, loss: 0.0007008292595855892  
step: 140, loss: 0.009974958375096321  
step: 150, loss: 0.0015244085807353258  
step: 160, loss: 0.0005462731933221221  
step: 170, loss: 0.005456698592752218  
step: 180, loss: 0.0055200280621647835  
step: 190, loss: 0.0026922558899968863  
step: 200, loss: 0.0012768697924911976  
step: 210, loss: 0.00862986221909523  
step: 220, loss: 0.004633184988051653  
step: 230, loss: 0.014645734801888466  
step: 240, loss: 0.02732672542333603  
step: 250, loss: 0.000674861017614603  
step: 260, loss: 7.915558671811596e-05  
step: 270, loss: 0.07209087908267975  
step: 280, loss: 0.010565550997853279  
step: 290, loss: 0.00041799465543590486  
step: 300, loss: 0.00034560143831185997  
step: 310, loss: 0.0013385117053985596  
step: 320, loss: 0.0003686240524984896  
step: 330, loss: 0.0012156200828030705  
step: 340, loss: 0.04583190008997917  
step: 350, loss: 0.007792208809405565  
step: 360, loss: 0.017232613638043404  
step: 370, loss: 0.0016458852915093303  
step: 380, loss: 0.00030639333999715745  
step: 390, loss: 0.020924609154462814  
step: 400, loss: 0.007975940592586994  
step: 410, loss: 0.001230638474225998  
step: 420, loss: 0.0065229786559939384
```

```
step: 430, loss: 0.0013623529812321067
step: 440, loss: 0.0004985319683328271
step: 450, loss: 0.0001880358177004382
step: 460, loss: 0.0002450074243824929
step: 470, loss: 0.006019569933414459
step: 480, loss: 0.03543807938694954
step: 490, loss: 0.00461282255128026
step: 500, loss: 0.0013275971869006753
step: 510, loss: 0.0033744059037417173
step: 520, loss: 0.001808775239624083
step: 530, loss: 0.004546301439404488
/usr/local/lib/python3.7/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=0.9426987060998151, f1=0.9456372772955687,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004104197956621647
step: 10, loss: 0.00013260514242574573
step: 20, loss: 0.0001568083098391071
step: 30, loss: 0.0002355258766328916
step: 40, loss: 0.07024520635604858
step: 50, loss: 0.004852311220020056
step: 60, loss: 0.016971848905086517
step: 70, loss: 0.00010283784649800509
step: 80, loss: 0.001005072146654129
step: 90, loss: 0.018203848972916603
step: 100, loss: 0.00040432284004054964
step: 110, loss: 0.0005428637377917767
step: 120, loss: 0.17475402355194092
step: 130, loss: 0.005426270887255669
step: 140, loss: 0.00014670705422759056
step: 150, loss: 0.005396115593612194
step: 160, loss: 6.174173904582858e-05
step: 170, loss: 8.009930024854839e-05
step: 180, loss: 0.0003006966144312173
step: 190, loss: 0.002133888192474842
step: 200, loss: 0.0003538133460097015
step: 210, loss: 0.000995305716060102
step: 220, loss: 0.0044607617892324924
step: 230, loss: 0.030588800087571144
step: 240, loss: 0.004693250171840191
step: 250, loss: 0.0009184451191686094
step: 260, loss: 0.0006834355881437659
step: 270, loss: 0.012400585226714611
step: 280, loss: 0.008917571976780891
```

```

step: 290, loss: 0.00018262697267346084
step: 300, loss: 0.010339714586734772
step: 310, loss: 0.00031652278266847134
step: 320, loss: 0.0013876657467335463
step: 330, loss: 0.015292257070541382
step: 340, loss: 0.030128994956612587
step: 350, loss: 0.017335470765829086
step: 360, loss: 0.002277444116771221
step: 370, loss: 0.0001254752278327942
step: 380, loss: 0.0011238360311836004
step: 390, loss: 0.0007911389693617821
step: 400, loss: 0.022163230925798416
step: 410, loss: 0.002084146486595273
step: 420, loss: 0.004912314936518669
step: 430, loss: 0.0003670294245239347
step: 440, loss: 0.0021832084748893976
step: 450, loss: 0.0008745152736082673
step: 460, loss: 0.005118262954056263
step: 470, loss: 0.06126585230231285
step: 480, loss: 0.012595158070325851
step: 490, loss: 0.0019889308605343103
step: 500, loss: 0.005663634277880192
step: 510, loss: 0.006605436094105244
step: 520, loss: 0.05207972973585129
step: 530, loss: 0.006590896286070347
/usr/local/lib/python3.7/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=0.9476609541454377, f1=0.9425925925925925,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001421018096152693
step: 10, loss: 0.027981096878647804
step: 20, loss: 4.064060340169817e-05
step: 30, loss: 0.0004211787017993629
step: 40, loss: 3.635768371168524e-05
step: 50, loss: 0.0031648105941712856
step: 60, loss: 0.001916629495099187
step: 70, loss: 6.843500159448013e-05
step: 80, loss: 0.0007523777894675732
step: 90, loss: 0.0008286657393909991
step: 100, loss: 0.00011161369911860675
step: 110, loss: 6.593576108571142e-05
step: 120, loss: 9.567713277647272e-05
step: 130, loss: 0.0045027039013803005
step: 140, loss: 0.006381685379892588

```

```

step: 150, loss: 5.3155625209910795e-05
step: 160, loss: 0.0016922337235882878
step: 170, loss: 0.0012684211833402514
step: 180, loss: 4.35837973782327e-05
step: 190, loss: 0.002699406584724784
step: 200, loss: 0.0012987845111638308
step: 210, loss: 0.0016162206884473562
step: 220, loss: 0.08806680887937546
step: 230, loss: 8.013381011551246e-05
step: 240, loss: 0.00024109453079290688
step: 250, loss: 0.00011997402907582
step: 260, loss: 0.004968826659023762
step: 270, loss: 0.01085593830794096
step: 280, loss: 0.026482924818992615
step: 290, loss: 7.646162703167647e-05
step: 300, loss: 0.009183881804347038
step: 310, loss: 0.00036003271816298366
step: 320, loss: 0.030368633568286896
step: 330, loss: 0.027771029621362686
step: 340, loss: 0.0039058958645910025
step: 350, loss: 0.0004306450136937201
step: 360, loss: 0.009286421351134777
step: 370, loss: 8.258757588919252e-05
step: 380, loss: 0.016290977597236633
step: 390, loss: 0.13608793914318085
step: 400, loss: 4.364004053059034e-05
step: 410, loss: 0.0007419977337121964
step: 420, loss: 0.0028243507258594036
step: 430, loss: 0.04617777466773987
step: 440, loss: 0.005107028875499964
step: 450, loss: 0.005264421459287405
step: 460, loss: 0.004203491378575563
step: 470, loss: 4.223546420689672e-05
step: 480, loss: 4.433971844264306e-05
step: 490, loss: 0.0007181463297456503
step: 500, loss: 0.00032106146682053804
step: 510, loss: 4.659069963963702e-05
step: 520, loss: 0.0029923520050942898
step: 530, loss: 6.901669985381886e-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=0.93989588263133, f1=0.941286989196806,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019877259619534016

```

step: 10, loss: 3.1921095796860754e-05
step: 20, loss: 0.00012892157246824354
step: 30, loss: 0.000224924020585604
step: 40, loss: 0.0024142060428857803
step: 50, loss: 2.6955232897307724e-05
step: 60, loss: 8.455403440166265e-05
step: 70, loss: 0.00012246092956047505
step: 80, loss: 0.006525801494717598
step: 90, loss: 4.639913095161319e-05
step: 100, loss: 3.591188578866422e-05
step: 110, loss: 2.8177139029139653e-05
step: 120, loss: 6.696554191876203e-05
step: 130, loss: 0.02320740930736065
step: 140, loss: 7.973586616571993e-05
step: 150, loss: 0.005205190274864435
step: 160, loss: 0.002007958712056279
step: 170, loss: 0.02462064102292061
step: 180, loss: 2.026144829869736e-05
step: 190, loss: 2.545772986195516e-05
step: 200, loss: 5.966426761006005e-05
step: 210, loss: 5.164538015378639e-05
step: 220, loss: 7.502169319195673e-05
step: 230, loss: 2.2668022211291827e-05
step: 240, loss: 6.540528556797653e-05
step: 250, loss: 0.0006255288608372211
step: 260, loss: 0.010446577332913876
step: 270, loss: 1.8048802303383127e-05
step: 280, loss: 0.00645392294973135
step: 290, loss: 0.040893204510211945
step: 300, loss: 2.2511569113703445e-05
step: 310, loss: 0.03763986751437187
step: 320, loss: 0.00013050706184003502
step: 330, loss: 2.4798857339192182e-05
step: 340, loss: 0.00012830545892938972
step: 350, loss: 0.003809631336480379
step: 360, loss: 2.0842664525844157e-05
step: 370, loss: 0.004020539578050375
step: 380, loss: 3.6624045606004074e-05
step: 390, loss: 0.0031260193791240454
step: 400, loss: 0.0001702926674624905
step: 410, loss: 3.259728691773489e-05
step: 420, loss: 0.00026810841518454254
step: 430, loss: 0.00016649758617859334
step: 440, loss: 0.001731481053866446
step: 450, loss: 7.069519051583484e-05
step: 460, loss: 0.0003902432508766651
step: 470, loss: 1.4796695722907316e-05
step: 480, loss: 6.55834810459055e-05

```

step: 490, loss: 0.0015144641511142254
step: 500, loss: 0.002266146009787917
step: 510, loss: 0.00022061661002226174
step: 520, loss: 0.00023508913000114262
step: 530, loss: 8.209343650378287e-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=0.9419600380589915, f1=0.9430051813471502,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030113516841083765
step: 10, loss: 4.146651554037817e-05
step: 20, loss: 2.5025272407219745e-05
step: 30, loss: 2.2656760847894475e-05
step: 40, loss: 5.342043732525781e-05
step: 50, loss: 0.0001327848294749856
step: 60, loss: 3.281072713434696e-05
step: 70, loss: 0.0004542506067082286
step: 80, loss: 2.2831778551335447e-05
step: 90, loss: 3.0441464332398027e-05
step: 100, loss: 4.106092092115432e-05
step: 110, loss: 1.7981690689339302e-05
step: 120, loss: 7.35687863198109e-05
step: 130, loss: 3.299918171251193e-05
step: 140, loss: 0.015658367425203323
step: 150, loss: 0.0004287780320737511
step: 160, loss: 0.0031979235354810953
step: 170, loss: 3.0445571610471234e-05
step: 180, loss: 2.4053546439972706e-05
step: 190, loss: 2.6780031475936994e-05
step: 200, loss: 0.0002841156965587288
step: 210, loss: 0.0002786662371363491
step: 220, loss: 8.719497418496758e-05
step: 230, loss: 0.0071058934554457664
step: 240, loss: 8.129162597469985e-05
step: 250, loss: 2.509643672965467e-05
step: 260, loss: 0.0022230034228414297
step: 270, loss: 0.0010962330270558596
step: 280, loss: 2.68735020654276e-05
step: 290, loss: 4.8777583288028836e-05
step: 300, loss: 2.684317405510228e-05
step: 310, loss: 2.3680961021455005e-05
step: 320, loss: 0.001359958085231483
step: 330, loss: 0.00013260073319543153
step: 340, loss: 0.009227894246578217

```



```

step: 350, loss: 0.0002733298751991242
step: 360, loss: 0.00012160040932940319
step: 370, loss: 0.00199871719814837
step: 380, loss: 0.1902458816766739
step: 390, loss: 0.0006349797477014363
step: 400, loss: 0.0018864895682781935
step: 410, loss: 0.0017420256044715643
step: 420, loss: 0.0007745864568278193
step: 430, loss: 1.3191143807489425e-05
step: 440, loss: 0.0004890434793196619
step: 450, loss: 3.418148116907105e-05
step: 460, loss: 2.319659870408941e-05
step: 470, loss: 4.0576553146820515e-05
step: 480, loss: 0.00022923940559849143
step: 490, loss: 2.6805962988873944e-05
step: 500, loss: 0.0007170645403675735
step: 510, loss: 1.5042566701595206e-05
step: 520, loss: 3.4017455618595704e-05
step: 530, loss: 0.0013075697934255004
/usr/local/lib/python3.7/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=0.9422098936662043, f1=0.9480937069361507,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.38048106036149e-05
step: 10, loss: 0.00010798553557833657
step: 20, loss: 4.668669862439856e-05
step: 30, loss: 0.00019623602565843612
step: 40, loss: 3.8872327422723174e-05
step: 50, loss: 2.505167867639102e-05
step: 60, loss: 0.017489222809672356
step: 70, loss: 2.2112617443781346e-05
step: 80, loss: 1.826473089749925e-05
step: 90, loss: 0.0006873502279631793
step: 100, loss: 0.0006033229292370379
step: 110, loss: 2.897441663662903e-05
step: 120, loss: 2.702091478568036e-05
step: 130, loss: 1.6733756638132036e-05
step: 140, loss: 0.002686602296307683
step: 150, loss: 2.4223478249041364e-05
step: 160, loss: 1.5917938071652316e-05
step: 170, loss: 2.4373663109145127e-05
step: 180, loss: 1.784760934242513e-05
step: 190, loss: 0.002596414415165782
step: 200, loss: 0.004793390166014433

```

```

step: 210, loss: 5.160149885341525e-05
step: 220, loss: 0.002388988621532917
step: 230, loss: 1.695348328212276e-05
step: 240, loss: 8.051384065765887e-05
step: 250, loss: 1.3869135727873072e-05
step: 260, loss: 1.598130438651424e-05
step: 270, loss: 2.050362127192784e-05
step: 280, loss: 5.723510912503116e-05
step: 290, loss: 0.00023164214508142322
step: 300, loss: 9.60757170105353e-05
step: 310, loss: 9.254027099814266e-05
step: 320, loss: 7.119979272829369e-05
step: 330, loss: 9.235645120497793e-05
step: 340, loss: 3.165204179822467e-05
step: 350, loss: 4.151822213316336e-05
step: 360, loss: 1.9963405065936968e-05
step: 370, loss: 8.165584586095065e-05
step: 380, loss: 0.00011335947056068107
step: 390, loss: 2.501818016753532e-05
step: 400, loss: 0.00018391915364190936
step: 410, loss: 0.004846268333494663
step: 420, loss: 0.00030984554905444384
step: 430, loss: 2.9383527362369932e-05
step: 440, loss: 6.972195842536166e-05
step: 450, loss: 4.414268187247217e-05
step: 460, loss: 0.00036282281507737935
step: 470, loss: 0.03681304678320885
step: 480, loss: 0.001691832672804594
step: 490, loss: 4.6779208787484095e-05
step: 500, loss: 0.00017196088447235525
step: 510, loss: 0.0003318168455734849
step: 520, loss: 5.896152288187295e-05
step: 530, loss: 4.858631655224599e-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=0.9418386491557222, f1=0.9392523364485983,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001482768275309354
step: 10, loss: 4.494928361964412e-05
step: 20, loss: 0.0003080111346207559
step: 30, loss: 0.0016116827027872205
step: 40, loss: 0.0009820363484323025
step: 50, loss: 0.02625558152794838
step: 60, loss: 0.0013357634888961911

```

step: 70, loss: 0.00010499197378521785
step: 80, loss: 3.95749302697368e-05
step: 90, loss: 4.42293421656359e-05
step: 100, loss: 3.4695804060902447e-05
step: 110, loss: 3.467306669335812e-05
step: 120, loss: 0.00012705705012194812
step: 130, loss: 0.0011615806724876165
step: 140, loss: 6.121994374552742e-05
step: 150, loss: 1.6435764337074943e-05
step: 160, loss: 1.994858575926628e-05
step: 170, loss: 4.352543328423053e-05
step: 180, loss: 0.00026713532861322165
step: 190, loss: 9.538896847516298e-05
step: 200, loss: 4.151897883275524e-05
step: 210, loss: 5.9739330026786774e-05
step: 220, loss: 8.755926683079451e-05
step: 230, loss: 2.32787224376807e-05
step: 240, loss: 0.0002679156023077667
step: 250, loss: 3.263916732976213e-05
step: 260, loss: 3.836033647530712e-05
step: 270, loss: 7.249271584441885e-05
step: 280, loss: 0.0015762023394927382
step: 290, loss: 0.00011121651186840609
step: 300, loss: 0.0012163042556494474
step: 310, loss: 6.14270320511423e-05
step: 320, loss: 2.4258368284790777e-05
step: 330, loss: 6.159176700748503e-05
step: 340, loss: 3.2631931389914826e-05
step: 350, loss: 0.00011818486382253468
step: 360, loss: 7.9195415310096e-05
step: 370, loss: 5.2266659622546285e-05
step: 380, loss: 0.00037252268521115184
step: 390, loss: 0.0002554641105234623
step: 400, loss: 6.802457210142165e-05
step: 410, loss: 0.0009251736337319016
step: 420, loss: 0.025953711941838264
step: 430, loss: 0.01940787397325039
step: 440, loss: 7.487643597414717e-05
step: 450, loss: 0.0001511493173893541
step: 460, loss: 2.5069810362765566e-05
step: 470, loss: 3.426471084821969e-05
step: 480, loss: 7.115917833289132e-05
step: 490, loss: 0.0008446328574791551
step: 500, loss: 0.0001415002770954743
step: 510, loss: 0.0011598714627325535
step: 520, loss: 2.919405778811779e-05
step: 530, loss: 3.47668283211533e-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=0.9448893075836081, f1=0.9429373246024322,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019697872921824455
step: 10, loss: 3.920426024706103e-05
step: 20, loss: 0.0004562213143799454
step: 30, loss: 2.817351014527958e-05
step: 40, loss: 2.6891795641859062e-05
step: 50, loss: 0.00033471279311925173
step: 60, loss: 1.9885123037965968e-05
step: 70, loss: 9.880581637844443e-05
step: 80, loss: 0.00015120007446967065
step: 90, loss: 0.02475050650537014
step: 100, loss: 0.0012163621140643954
step: 110, loss: 2.04886728170095e-05
step: 120, loss: 9.137719462160021e-05
step: 130, loss: 2.2623215045314282e-05
step: 140, loss: 0.00012757470540236682
step: 150, loss: 0.00022826433996669948
step: 160, loss: 0.001828933833166957
step: 170, loss: 0.00019162491662427783
step: 180, loss: 0.00017418622155673802
step: 190, loss: 1.9803304894594476e-05
step: 200, loss: 2.54834012594074e-05
step: 210, loss: 2.8999569622101262e-05
step: 220, loss: 1.8104527043760754e-05
step: 230, loss: 2.2920965420780703e-05
step: 240, loss: 2.372195740463212e-05
step: 250, loss: 2.0224093532306142e-05
step: 260, loss: 2.099127414112445e-05
step: 270, loss: 1.974707265617326e-05
step: 280, loss: 5.437879008240998e-05
step: 290, loss: 2.527082870074082e-05
step: 300, loss: 3.1539995688945055e-05
step: 310, loss: 0.00040885983617044985
step: 320, loss: 0.0030905422754585743
step: 330, loss: 0.0013285009190440178
step: 340, loss: 6.254420441109687e-05
step: 350, loss: 1.789972884580493e-05
step: 360, loss: 1.887946746137459e-05
step: 370, loss: 5.4234038543654606e-05
step: 380, loss: 2.8086906240787357e-05
step: 390, loss: 0.25512778759002686
step: 400, loss: 0.00015404814621433616
```

```

step: 410, loss: 0.0007532477611675858
step: 420, loss: 2.4399794710916467e-05
step: 430, loss: 0.006370437331497669
step: 440, loss: 0.00012014888488920406
step: 450, loss: 0.00012263283133506775
step: 460, loss: 0.0011560063576325774
step: 470, loss: 3.263206235715188e-05
step: 480, loss: 8.133116352837533e-05
step: 490, loss: 7.029345579212531e-05
step: 500, loss: 0.0002906704612541944
step: 510, loss: 0.00012583429634105414
step: 520, loss: 0.00014222414756659418
step: 530, loss: 0.0004656133824028075
/usr/local/lib/python3.7/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=0.9451865499769692, f1=0.9437070938215102,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019256504019722342
step: 10, loss: 1.843595237005502e-05
step: 20, loss: 0.0035542957484722137
step: 30, loss: 4.609670941135846e-05
step: 40, loss: 0.0020406432449817657
step: 50, loss: 0.00030876637902110815
step: 60, loss: 1.8562634068075567e-05
step: 70, loss: 4.277572224964388e-05
step: 80, loss: 5.4031857871450484e-05
step: 90, loss: 3.389078119653277e-05
step: 100, loss: 0.00016033777501434088
step: 110, loss: 2.457510527165141e-05
step: 120, loss: 5.343140219338238e-05
step: 130, loss: 0.021955495700240135
step: 140, loss: 0.0016670410986989737
step: 150, loss: 1.8197584722656757e-05
step: 160, loss: 0.0002931642811745405
step: 170, loss: 0.0008799899951554835
step: 180, loss: 2.001543361984659e-05
step: 190, loss: 0.0018315041670575738
step: 200, loss: 0.00017476867651566863
step: 210, loss: 2.8497457606135868e-05
step: 220, loss: 2.8959131668671034e-05
step: 230, loss: 4.412557609612122e-05
step: 240, loss: 2.459340976201929e-05
step: 250, loss: 3.651829683803953e-05
step: 260, loss: 0.0007044658996164799

```

```

step: 270, loss: 2.0916861103614792e-05
step: 280, loss: 1.969132608792279e-05
step: 290, loss: 0.0015768918674439192
step: 300, loss: 0.00017605346511118114
step: 310, loss: 0.00027111551025882363
step: 320, loss: 0.0019686538726091385
step: 330, loss: 0.0002940593403764069
step: 340, loss: 8.821509254630655e-05
step: 350, loss: 6.761283293599263e-05
step: 360, loss: 7.459038170054555e-05
step: 370, loss: 2.3066506400937214e-05
step: 380, loss: 5.847620195709169e-05
step: 390, loss: 0.02532777190208435
step: 400, loss: 0.0001652916253078729
step: 410, loss: 2.4741982997511514e-05
step: 420, loss: 0.00022370461374521255
step: 430, loss: 7.217629172373563e-05
step: 440, loss: 9.508922812528908e-05
step: 450, loss: 0.00013843107444699854
step: 460, loss: 0.0011235822457820177
step: 470, loss: 1.8979953892994672e-05
step: 480, loss: 5.088153193355538e-05
step: 490, loss: 0.00022762760636396706
step: 500, loss: 2.7117001081933267e-05
step: 510, loss: 0.00033723359229043126
step: 520, loss: 0.00130398478358984
step: 530, loss: 0.0020984390284866095
/usr/local/lib/python3.7/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=0.9415067852129153, f1=0.9434843531060252,
best_f1=0.945387792565397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3155422240961343e-05
step: 10, loss: 9.76775772869587e-05
step: 20, loss: 0.000288078241283074
step: 30, loss: 8.065911970334128e-05
step: 40, loss: 0.07816372811794281
step: 50, loss: 0.0007507515256293118
step: 60, loss: 1.5429664927069098e-05
step: 70, loss: 0.00039480184204876423
step: 80, loss: 8.093736687442288e-05
step: 90, loss: 2.377385135332588e-05
step: 100, loss: 0.0005262094782665372
step: 110, loss: 0.0014966438757255673
step: 120, loss: 0.00026237056590616703

```

```

step: 130, loss: 0.09735741466283798
step: 140, loss: 2.195989691244904e-05
step: 150, loss: 3.810011548921466e-05
step: 160, loss: 0.00030073378002271056
step: 170, loss: 0.0002963572624139488
step: 180, loss: 2.5836068743956275e-05
step: 190, loss: 4.216632805764675e-05
step: 200, loss: 7.081469084369019e-05
step: 210, loss: 0.0027792309410870075
step: 220, loss: 1.956459709617775e-05
step: 230, loss: 0.0008992250077426434
step: 240, loss: 2.837761894625146e-05
step: 250, loss: 0.00010986263077938929
step: 260, loss: 4.8736317694419995e-05
step: 270, loss: 2.694772047107108e-05
step: 280, loss: 2.859736923710443e-05
step: 290, loss: 1.5694402463850565e-05
step: 300, loss: 1.969863660633564e-05
step: 310, loss: 8.663991320645437e-05
step: 320, loss: 2.7973886972176842e-05
step: 330, loss: 3.7518537283176556e-05
step: 340, loss: 1.9404453269089572e-05
step: 350, loss: 2.3333870558417402e-05
step: 360, loss: 0.0016234290087595582
step: 370, loss: 1.4643826034443919e-05
step: 380, loss: 0.0003486373752821237
step: 390, loss: 0.00024123738694470376
step: 400, loss: 2.3956443328643218e-05
step: 410, loss: 3.725615897565149e-05
step: 420, loss: 2.2347452613757923e-05
step: 430, loss: 0.0004296850529499352
step: 440, loss: 3.9312042645178735e-05
step: 450, loss: 5.178979699849151e-05
step: 460, loss: 2.659756864886731e-05
step: 470, loss: 0.0025841251481324434
step: 480, loss: 1.4658745385531802e-05
step: 490, loss: 1.791830982256215e-05
step: 500, loss: 0.0003734048514161259
step: 510, loss: 0.00010792462853714824
step: 520, loss: 1.709504795144312e-05
step: 530, loss: 0.0005166491027921438
/usr/local/lib/python3.7/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=0.9447565543071161, f1=0.9434843531060252,
best_f1=0.945387792565397
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32

```

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

5742it [00:23, 244.28it/s]

load_f1 = 0.9463459759481962

real_f1 = 0.9458082445576655

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, 244.08it/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: 0.576972484588623
step: 10, loss: 0.41053372621536255
step: 20, loss: 0.3849208354949951
step: 30, loss: 0.31167978048324585
step: 40, loss: 0.18829679489135742
step: 50, loss: 0.4123140275478363
step: 60, loss: 0.27881836891174316
step: 70, loss: 0.16589713096618652
step: 80, loss: 0.20385494828224182
step: 90, loss: 0.2774759531021118
step: 100, loss: 0.3535251319408417
step: 110, loss: 0.17177064716815948
step: 120, loss: 0.14390145242214203
step: 130, loss: 0.1476907581090927
step: 140, loss: 0.22061105072498322
step: 150, loss: 0.21194739639759064
step: 160, loss: 0.25725409388542175
step: 170, loss: 0.2763320505619049
step: 180, loss: 0.08756178617477417
step: 190, loss: 0.25437918305397034
step: 200, loss: 0.2385900467634201
step: 210, loss: 0.21625670790672302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6104868913857678, f1=0.6653620352250489,
best_f1=0.6653620352250489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08304663002490997

```

```

step: 10, loss: 0.17867940664291382
step: 20, loss: 0.19401052594184875
step: 30, loss: 0.22106300294399261
step: 40, loss: 0.1377921849489212
step: 50, loss: 0.12870623171329498
step: 60, loss: 0.3473048210144043
step: 70, loss: 0.09565117955207825
step: 80, loss: 0.21209344267845154
step: 90, loss: 0.06940939277410507
step: 100, loss: 0.009896154515445232
step: 110, loss: 0.11108572036027908
step: 120, loss: 0.13263718783855438
step: 130, loss: 0.014410600066184998
step: 140, loss: 0.17316095530986786
step: 150, loss: 0.11617239564657211
step: 160, loss: 0.20647385716438293
step: 170, loss: 0.12078174203634262
step: 180, loss: 0.1687774509191513
step: 190, loss: 0.21079374849796295
step: 200, loss: 0.03868618607521057
step: 210, loss: 0.1446363925933838
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.7347670250896058,
best_f1=0.7347670250896058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03286135196685791
step: 10, loss: 0.159610316157341
step: 20, loss: 0.0652281790971756
step: 30, loss: 0.14443816244602203
step: 40, loss: 0.12742814421653748
step: 50, loss: 0.10029033571481705
step: 60, loss: 0.10509466379880905
step: 70, loss: 0.05606663227081299
step: 80, loss: 0.054911550134420395
step: 90, loss: 0.005982725881040096
step: 100, loss: 0.2113213688135147
step: 110, loss: 0.18986158072948456
step: 120, loss: 0.09765984863042831
step: 130, loss: 0.13545943796634674
step: 140, loss: 0.07814542949199677

```

```

step: 150, loss: 0.2083221822977066
step: 160, loss: 0.006504957098513842
step: 170, loss: 0.05467325821518898
step: 180, loss: 0.1945190280675888
step: 190, loss: 0.19660544395446777
step: 200, loss: 0.035910338163375854
step: 210, loss: 0.07581975311040878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6771378708551483, f1=0.730837789661319,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10781216621398926
step: 10, loss: 0.10928212106227875
step: 20, loss: 0.030569477006793022
step: 30, loss: 0.15251567959785461
step: 40, loss: 0.006826837081462145
step: 50, loss: 0.15717941522598267
step: 60, loss: 0.10106342285871506
step: 70, loss: 0.10966095328330994
step: 80, loss: 0.03941686823964119
step: 90, loss: 0.012677529826760292
step: 100, loss: 0.21604317426681519
step: 110, loss: 0.10865623503923416
step: 120, loss: 0.08291685581207275
step: 130, loss: 0.13897427916526794
step: 140, loss: 0.050288647413253784
step: 150, loss: 0.0729653388261795
step: 160, loss: 0.028432078659534454
step: 170, loss: 0.09286188334226608
step: 180, loss: 0.34876522421836853
step: 190, loss: 0.08979079872369766
step: 200, loss: 0.0939415693283081
step: 210, loss: 0.11895468086004257
/usr/local/lib/python3.7/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=0.6433566433566433, f1=0.7067137809187279,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08047252893447876
step: 10, loss: 0.046038027852773666
step: 20, loss: 0.04995592311024666
step: 30, loss: 0.018243085592985153
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.03637795150279999
step: 50, loss: 0.05249720439314842
step: 60, loss: 0.08792957663536072
step: 70, loss: 0.04083247482776642
step: 80, loss: 0.02633778005838394
step: 90, loss: 0.06305944919586182
step: 100, loss: 0.003274078480899334
step: 110, loss: 0.13122332096099854
step: 120, loss: 0.0995638445019722
step: 130, loss: 0.021159838885068893
step: 140, loss: 0.021696215495467186
step: 150, loss: 0.042235083878040314
step: 160, loss: 0.2247985601425171
step: 170, loss: 0.0563262403011322
step: 180, loss: 0.07823428511619568
step: 190, loss: 0.035735078155994415
step: 200, loss: 0.08113216608762741
step: 210, loss: 0.007208531256765127
/usr/local/lib/python3.7/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=0.6567164179104479, f1=0.725233644859813,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011800428852438927
step: 10, loss: 0.11861511319875717
step: 20, loss: 0.025342632085084915
step: 30, loss: 0.0026520031969994307
step: 40, loss: 0.03585263341665268
step: 50, loss: 0.0027584631461650133
step: 60, loss: 0.04391296207904816
step: 70, loss: 0.014543503522872925
step: 80, loss: 0.05408584326505661
step: 90, loss: 0.10371188819408417
step: 100, loss: 0.016757385805249214
step: 110, loss: 0.017305200919508934
step: 120, loss: 0.09467922896146774
step: 130, loss: 0.0454457588493824
step: 140, loss: 0.05576934292912483
step: 150, loss: 0.0312621183693409

```

```

step: 160, loss: 0.015092654153704643
step: 170, loss: 0.09174612909555435
step: 180, loss: 0.01613743230700493
step: 190, loss: 0.09447024762630463
step: 200, loss: 0.03301168978214264
step: 210, loss: 0.032931774854660034
/usr/local/lib/python3.7/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=0.6467065868263474, f1=0.7243460764587525,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10744298994541168
step: 10, loss: 0.0060058641247451305
step: 20, loss: 0.00676707411184907
step: 30, loss: 0.06795086711645126
step: 40, loss: 0.029780695214867592
step: 50, loss: 0.04088051989674568
step: 60, loss: 0.009306696243584156
step: 70, loss: 0.01373467966914177
step: 80, loss: 0.031091108918190002
step: 90, loss: 0.08060738444328308
step: 100, loss: 0.004026887938380241
step: 110, loss: 0.044681429862976074
step: 120, loss: 0.1009393185377121
step: 130, loss: 0.016305018216371536
step: 140, loss: 0.006617147941142321
step: 150, loss: 0.017232144251465797
step: 160, loss: 0.02361966110765934
step: 170, loss: 0.0004290569922886789
step: 180, loss: 0.10842688381671906
step: 190, loss: 0.0674242451786995
step: 200, loss: 0.005464911460876465
step: 210, loss: 0.016733268275856972
/usr/local/lib/python3.7/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=0.6573146292585169, f1=0.7006109979633403,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06923745572566986
step: 10, loss: 0.040614813566207886
step: 20, loss: 0.0010629701428115368
step: 30, loss: 0.10232435911893845

```

```

step: 40, loss: 0.00238005630671978
step: 50, loss: 0.0010787678183987737
step: 60, loss: 0.005645532160997391
step: 70, loss: 0.12271589785814285
step: 80, loss: 0.03210937976837158
step: 90, loss: 0.02017362415790558
step: 100, loss: 0.07921559363603592
step: 110, loss: 0.012161068618297577
step: 120, loss: 0.0019764634780585766
step: 130, loss: 0.0008292788406834006
step: 140, loss: 0.06480303406715393
step: 150, loss: 0.008793405257165432
step: 160, loss: 0.0350445881485939
step: 170, loss: 0.0012627181131392717
step: 180, loss: 0.030131269246339798
step: 190, loss: 0.034424539655447006
step: 200, loss: 0.011761009693145752
step: 210, loss: 0.06100578233599663
/usr/local/lib/python3.7/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=0.6639839034205232, f1=0.7119341563786008,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015273146564140916
step: 10, loss: 0.011218369007110596
step: 20, loss: 0.015086856670677662
step: 30, loss: 0.01868281699717045
step: 40, loss: 0.02024540677666664
step: 50, loss: 0.08820515871047974
step: 60, loss: 0.029743971303105354
step: 70, loss: 0.0018745906418189406
step: 80, loss: 0.002144751138985157
step: 90, loss: 0.0004676024545915425
step: 100, loss: 0.002721830504015088
step: 110, loss: 0.030184432864189148
step: 120, loss: 0.0011842579115182161
step: 130, loss: 0.008853467181324959
step: 140, loss: 0.006072746589779854
step: 150, loss: 0.06188811734318733
step: 160, loss: 0.001245402731001377
step: 170, loss: 0.00561012327671051
step: 180, loss: 0.030640581622719765
step: 190, loss: 0.000664268562104553
step: 200, loss: 0.16869190335273743
step: 210, loss: 0.008619902655482292

```

```

/usr/local/lib/python3.7/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=0.6413502109704642, f1=0.6984815618221258,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032123082783073187
step: 10, loss: 0.01159152016043663
step: 20, loss: 0.0005412348546087742
step: 30, loss: 0.19283854961395264
step: 40, loss: 0.0003252193273510784
step: 50, loss: 0.007107475306838751
step: 60, loss: 0.027200376614928246
step: 70, loss: 0.1076495349407196
step: 80, loss: 0.012038943357765675
step: 90, loss: 0.27139970660209656
step: 100, loss: 0.002666895044967532
step: 110, loss: 0.0004829354875255376
step: 120, loss: 0.013389146886765957
step: 130, loss: 0.022336123511195183
step: 140, loss: 0.002755995374172926
step: 150, loss: 0.0710543543100357
step: 160, loss: 0.018184930086135864
step: 170, loss: 0.0011029463494196534
step: 180, loss: 0.007987617515027523
step: 190, loss: 0.056998360902071
step: 200, loss: 0.0004674251249525696
step: 210, loss: 0.05773036926984787
/usr/local/lib/python3.7/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=0.640973630831643, f1=0.6995884773662552,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013422622345387936
step: 10, loss: 0.029533078894019127
step: 20, loss: 0.01707177609205246
step: 30, loss: 0.00046507539809681475
step: 40, loss: 0.01025173719972372
step: 50, loss: 0.013399047777056694
step: 60, loss: 0.033712584525346756
step: 70, loss: 0.03427805379033089
step: 80, loss: 0.03137849271297455
step: 90, loss: 0.000651761656627059

```



```

step: 100, loss: 0.024004893377423286
step: 110, loss: 0.013131178915500641
step: 120, loss: 0.013336632400751114
step: 130, loss: 0.026580216363072395
step: 140, loss: 0.0002448550658300519
step: 150, loss: 0.0006817918037995696
step: 160, loss: 0.027942148968577385
step: 170, loss: 0.0035073338076472282
step: 180, loss: 0.0010650415206328034
step: 190, loss: 0.000287192000541836
step: 200, loss: 0.00044974873890168965
step: 210, loss: 0.0031074390280991793
/usr/local/lib/python3.7/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=0.6442105263157895, f1=0.7053763440860215,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02783452346920967
step: 10, loss: 0.0001745164190651849
step: 20, loss: 0.039618633687496185
step: 30, loss: 0.0003358133544679731
step: 40, loss: 0.014467405155301094
step: 50, loss: 0.007978443056344986
step: 60, loss: 0.004260076675564051
step: 70, loss: 0.03604035824537277
step: 80, loss: 0.012440981343388557
step: 90, loss: 0.023517776280641556
step: 100, loss: 0.01965416967868805
step: 110, loss: 0.0033783356193453074
step: 120, loss: 0.015408353880047798
step: 130, loss: 0.0012433371739462018
step: 140, loss: 0.0003603199729695916
step: 150, loss: 0.009508916176855564
step: 160, loss: 0.0003185684035997838
step: 170, loss: 0.0018364283023402095
step: 180, loss: 0.0006857654079794884
step: 190, loss: 0.004047839902341366
step: 200, loss: 0.00043225105036981404
step: 210, loss: 0.00413315836340189
/usr/local/lib/python3.7/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=0.6376811594202899, f1=0.7068965517241379,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051053211092948914
step: 10, loss: 0.00012779200915247202
step: 20, loss: 0.016401134431362152
step: 30, loss: 0.05930207669734955
step: 40, loss: 0.01775183156132698
step: 50, loss: 0.000693206035066396
step: 60, loss: 0.0006294460035860538
step: 70, loss: 0.04888704791665077
step: 80, loss: 0.004196947440505028
step: 90, loss: 0.0012258993228897452
step: 100, loss: 0.00038164283614605665
step: 110, loss: 0.0007677879184484482
step: 120, loss: 0.00016709076589904726
step: 130, loss: 0.00023532539489679039
step: 140, loss: 0.01995137333869934
step: 150, loss: 0.0004354367556516081
step: 160, loss: 0.02859356440603733
step: 170, loss: 0.0025056509766727686
step: 180, loss: 0.03173825144767761
step: 190, loss: 0.0005848247092217207
step: 200, loss: 0.0007258767727762461
step: 210, loss: 0.016458654776215553
/usr/local/lib/python3.7/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=0.6519114688128774, f1=0.7166666666666667,
best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027603405760601163
step: 10, loss: 0.03578241169452667
step: 20, loss: 0.0015729541191831231
step: 30, loss: 0.0003981462796218693
step: 40, loss: 0.00033436386729590595
step: 50, loss: 0.0010228741448372602
step: 60, loss: 0.00564956571906805
step: 70, loss: 0.0019744865130633116
step: 80, loss: 0.0003432270896155387
step: 90, loss: 0.006057647988200188
step: 100, loss: 0.0009261500672437251
step: 110, loss: 0.0003516579745337367
step: 120, loss: 0.019709967076778412
step: 130, loss: 0.010446174070239067
step: 140, loss: 0.004068680573254824
step: 150, loss: 0.000847190385684371

```

```

step: 160, loss: 0.003369442420080304
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.0003050489758607
step: 180, loss: 0.0005221815663389862
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.011245028115808964
step: 200, loss: 0.008269392885267735
step: 210, loss: 0.0004146217543166131
/usr/local/lib/python3.7/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=0.6315789473684211, f1=0.7, best_f1=0.730837789661319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003138025349471718
step: 10, loss: 0.00013207348820287734
step: 20, loss: 0.0028296031523495913
step: 30, loss: 0.019936226308345795
step: 40, loss: 0.0003210308204870671
step: 50, loss: 0.00020376074826344848
step: 60, loss: 0.000505061645526439
step: 70, loss: 0.00014214077964425087
step: 80, loss: 0.0004983244580216706
step: 90, loss: 0.0170454028993845
step: 100, loss: 0.0002487664169166237
step: 110, loss: 0.0001683751615928486
step: 120, loss: 0.00023559603141620755
step: 130, loss: 0.00025813927641138434
step: 140, loss: 0.00010972926975227892
step: 150, loss: 0.0002091718779411167
step: 160, loss: 0.000574990117456764
step: 170, loss: 0.0002895425714086741
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 180, loss: 0.0001678124099271372
step: 190, loss: 0.00017247458163183182
step: 200, loss: 0.001529208617284894
step: 210, loss: 0.0002670694957487285
/usr/local/lib/python3.7/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=0.6260504201680671, f1=0.6956521739130435,
best_f1=0.730837789661319
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 509.83it/s]
load_f1 = 0.673469387755102
real_f1 = 0.6582809224318658
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:17, 256.75it/s]

```

3.1.5 Walmart-Amazon - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

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

step: 0, loss: 0.5710726380348206
step: 10, loss: 0.36774593591690063
step: 20, loss: 0.30501800775527954
step: 30, loss: 0.41471415758132935
step: 40, loss: 0.43674060702323914
step: 50, loss: 0.27824240922927856
step: 60, loss: 0.2496550977230072
step: 70, loss: 0.25966012477874756
step: 80, loss: 0.2183169722557068
step: 90, loss: 0.23465096950531006
step: 100, loss: 0.3275453448295593
step: 110, loss: 0.4850616753101349
step: 120, loss: 0.1220352053642273
step: 130, loss: 0.11263804882764816
step: 140, loss: 0.06510598957538605
step: 150, loss: 0.1532764881849289
step: 160, loss: 0.07036319375038147
step: 170, loss: 0.3261868953704834
step: 180, loss: 0.035373613238334656
step: 190, loss: 0.2235872894525528

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

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

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

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=0.7040816326530612, f1=0.7208121827411167,
best_f1=0.7208121827411167

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

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

step: 0, loss: 0.19229404628276825
step: 10, loss: 0.09011892974376678
step: 20, loss: 0.06545571237802505
step: 30, loss: 0.15874487161636353
step: 40, loss: 0.34184327721595764
step: 50, loss: 0.08190109580755234

```

step: 60, loss: 0.16120803356170654
step: 70, loss: 0.0948096513748169
step: 80, loss: 0.10128152370452881
step: 90, loss: 0.2301582545042038
step: 100, loss: 0.016530636698007584
step: 110, loss: 0.10541420429944992
step: 120, loss: 0.21329325437545776
step: 130, loss: 0.04748983681201935
step: 140, loss: 0.042657870799303055
step: 150, loss: 0.07224074006080627
step: 160, loss: 0.03415602818131447
step: 170, loss: 0.058796145021915436
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.1744668334722519
step: 190, loss: 0.03542976453900337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7909604519774011, f1=0.7814207650273224,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021156350150704384
step: 10, loss: 0.04746469110250473
step: 20, loss: 0.053812410682439804
step: 30, loss: 0.027723468840122223
step: 40, loss: 0.024918777868151665
step: 50, loss: 0.10247622430324554
step: 60, loss: 0.013097263872623444
step: 70, loss: 0.0822196826338768
step: 80, loss: 0.035860102623701096
step: 90, loss: 0.23300504684448242
step: 100, loss: 0.059049446135759354
step: 110, loss: 0.03907618671655655
step: 120, loss: 0.029941510409116745
step: 130, loss: 0.008989905007183552
step: 140, loss: 0.026491563767194748
step: 150, loss: 0.10440070182085037
step: 160, loss: 0.01885887421667576
step: 170, loss: 0.07337380945682526
step: 180, loss: 0.047619376331567764
step: 190, loss: 0.27420926094055176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.802030456852792, f1=0.8010204081632654,
best_f1=0.8010204081632654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011508631519973278
step: 10, loss: 0.09010881930589676
step: 20, loss: 0.030001316219568253
step: 30, loss: 0.003393118269741535
step: 40, loss: 0.009841321036219597
step: 50, loss: 0.010460563004016876
step: 60, loss: 0.0022385460324585438
step: 70, loss: 0.016748417168855667
step: 80, loss: 0.12167617678642273
step: 90, loss: 0.007027163170278072
step: 100, loss: 0.04963454604148865
step: 110, loss: 0.0018770830938592553
step: 120, loss: 0.013540047220885754
step: 130, loss: 0.307010293006897
step: 140, loss: 0.006051418837159872
step: 150, loss: 0.020849648863077164
step: 160, loss: 0.002274663420394063
step: 170, loss: 0.11351002752780914
step: 180, loss: 0.0014542165445163846
step: 190, loss: 0.22200563549995422
/usr/local/lib/python3.7/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=0.7724867724867726, f1=0.7634408602150539,
best_f1=0.8010204081632654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0475030392408371
step: 10, loss: 0.008696120232343674
step: 20, loss: 0.02259037271142006
step: 30, loss: 0.00245650508441031
step: 40, loss: 0.04294779151678085
step: 50, loss: 0.05214858800172806
step: 60, loss: 0.056271232664585114
step: 70, loss: 0.0009811443742364645
step: 80, loss: 0.037205763161182404

```



```

step: 90, loss: 0.007389598526060581
step: 100, loss: 0.00457768514752388
step: 110, loss: 0.003359634894877672
step: 120, loss: 0.011183470487594604
step: 130, loss: 0.045949846506118774
step: 140, loss: 0.08033318817615509
step: 150, loss: 0.05198505148291588
step: 160, loss: 0.003483239095658064
step: 170, loss: 0.014285658486187458
step: 180, loss: 0.008211602456867695
step: 190, loss: 0.01853170432150364
/usr/local/lib/python3.7/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=0.7979274611398963, f1=0.7989690721649483,
best_f1=0.8010204081632654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009662318974733353
step: 10, loss: 0.027937209233641624
step: 20, loss: 0.015866288915276527
step: 30, loss: 0.0626462921500206
step: 40, loss: 0.05438223108649254
step: 50, loss: 0.005805719643831253
step: 60, loss: 0.06500686705112457
step: 70, loss: 0.05351778119802475
step: 80, loss: 0.0036426701117306948
step: 90, loss: 0.0006694190087728202
step: 100, loss: 0.001894889515824616
step: 110, loss: 0.012737932614982128
step: 120, loss: 0.0008780842763371766
step: 130, loss: 0.0007869470282457769
step: 140, loss: 0.1553245335817337
step: 150, loss: 0.0040974486619234085
step: 160, loss: 0.038416117429733276
step: 170, loss: 0.028010722249746323
step: 180, loss: 0.005576267838478088
step: 190, loss: 0.005531102418899536
/usr/local/lib/python3.7/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=0.772972972972973, f1=0.8086253369272237,
best_f1=0.8010204081632654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009132416918873787

```

```

step: 10, loss: 0.029168931767344475
step: 20, loss: 0.08011867851018906
step: 30, loss: 0.009973143227398396
step: 40, loss: 0.0020084234420210123
step: 50, loss: 0.0014733860734850168
step: 60, loss: 0.0027252850122749805
step: 70, loss: 0.0059769973158836365
step: 80, loss: 0.0986361876130104
step: 90, loss: 0.0003427375922910869
step: 100, loss: 0.002030353993177414
step: 110, loss: 0.005165365058928728
step: 120, loss: 0.004428466781973839
step: 130, loss: 0.007385782431811094
step: 140, loss: 0.011937200091779232
step: 150, loss: 0.0008666929206810892
step: 160, loss: 0.04072452709078789
step: 170, loss: 0.0023152900394052267
step: 180, loss: 0.00194872310385108
step: 190, loss: 0.000781464739702642
/usr/local/lib/python3.7/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=0.7882037533512065, f1=0.786096256684492,
best_f1=0.8010204081632654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009692707099020481
step: 10, loss: 0.0035980220418423414
step: 20, loss: 0.011635993607342243
step: 30, loss: 0.014766023494303226
step: 40, loss: 0.0007278561824932694
step: 50, loss: 0.001116706756874919
step: 60, loss: 0.0005450972239486873
step: 70, loss: 0.0011445400305092335
step: 80, loss: 0.00011151580838486552
step: 90, loss: 0.0008385122055187821
step: 100, loss: 0.040702708065509796
step: 110, loss: 0.009776987135410309
step: 120, loss: 0.001757719088345766
step: 130, loss: 0.001936832326464355
step: 140, loss: 0.0016104986425489187
step: 150, loss: 0.0030931737273931503
step: 160, loss: 0.006375009194016457
step: 170, loss: 0.0013289585476741195
step: 180, loss: 0.0045919883996248245
step: 190, loss: 0.0011990313651040196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8044077134986225, f1=0.814404432132964,
best_f1=0.814404432132964
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004483606026042253
step: 10, loss: 0.021849557757377625
step: 20, loss: 0.003400747897103429
step: 30, loss: 0.0011761578498408198
step: 40, loss: 0.002720098476856947
step: 50, loss: 0.0010242852149531245
step: 60, loss: 0.00659537548199296
step: 70, loss: 0.0009117239969782531
step: 80, loss: 0.0008363056113012135
step: 90, loss: 0.015886448323726654
step: 100, loss: 0.04837403818964958
step: 110, loss: 0.0015081053134053946
step: 120, loss: 0.0731918141245842
step: 130, loss: 0.01606569066643715
step: 140, loss: 0.007424388080835342
step: 150, loss: 0.0024936769623309374
step: 160, loss: 0.001494289841502905
step: 170, loss: 0.02406379021704197
step: 180, loss: 0.0009182212525047362
step: 190, loss: 0.00021456459944602102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8130081300813009, f1=0.8075880758807588,
best_f1=0.8075880758807588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005688493838533759
step: 10, loss: 0.0003868693602271378
step: 20, loss: 0.018743623048067093
step: 30, loss: 0.00013581918028648943
step: 40, loss: 0.034261126071214676

```

```

step: 50, loss: 0.00029161045677028596
step: 60, loss: 0.00041367506491951644
step: 70, loss: 0.0005110210040584207
step: 80, loss: 0.0008227089419960976
step: 90, loss: 0.00018675083993002772
step: 100, loss: 0.0003385675372555852
step: 110, loss: 0.004082809668034315
step: 120, loss: 0.005990130361169577
step: 130, loss: 0.0013784762704744935
step: 140, loss: 0.0003312165499664843
step: 150, loss: 0.00017779746849555522
step: 160, loss: 0.000603700871579349
step: 170, loss: 0.00024235437740571797
step: 180, loss: 0.00017238316650036722
step: 190, loss: 0.000788882200140506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8, f1=0.7989130434782609, best_f1=0.8075880758807588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005438777734525502
step: 10, loss: 0.00025070333504118025
step: 20, loss: 0.054814063012599945
step: 30, loss: 0.025891872122883797
step: 40, loss: 0.0024334783665835857
step: 50, loss: 0.0010078259510919452
step: 60, loss: 9.031457739183679e-05
step: 70, loss: 0.00026765139773488045
step: 80, loss: 0.00011049964086851105
step: 90, loss: 7.62056079111062e-05
step: 100, loss: 0.0038353889249265194
step: 110, loss: 0.0002365531399846077
step: 120, loss: 0.00025711627677083015
step: 130, loss: 0.0003786565794143826
step: 140, loss: 0.000746081059332937
step: 150, loss: 0.00020702779875136912
step: 160, loss: 0.00021932265372015536
step: 170, loss: 0.00029817837639711797
step: 180, loss: 0.004030339419841766
step: 190, loss: 0.00024472433142364025
/usr/local/lib/python3.7/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=0.7928388746803069, f1=0.805128205128205,
best_f1=0.8075880758807588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.00016661420522723347
step: 10, loss: 0.007091882172971964
step: 20, loss: 0.00012621651694644243
step: 30, loss: 0.00043306947918608785
step: 40, loss: 0.00034835864789783955
step: 50, loss: 0.00030649182735942304
step: 60, loss: 0.0006771246553398669
step: 70, loss: 9.075440175365657e-05
step: 80, loss: 0.00015043573512230068
step: 90, loss: 0.00013578662765212357
step: 100, loss: 0.00024028470215853304
step: 110, loss: 0.00018673503655008972
step: 120, loss: 0.00021037834812887013
step: 130, loss: 0.00017550186021253467
step: 140, loss: 0.00024763410328887403
step: 150, loss: 0.00010666877642506734
step: 160, loss: 4.93510378873907e-05
step: 170, loss: 0.00032518833177164197
step: 180, loss: 5.8322846598457545e-05
step: 190, loss: 0.00010731421934906393
```

```
/usr/local/lib/python3.7/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=0.8054794520547945, f1=0.7956989247311829,
best_f1=0.8075880758807588
```

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

An input tensor was not cuda.

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

```
step: 0, loss: 0.0003165063390042633
step: 10, loss: 0.00010151842434424907
step: 20, loss: 0.0004981958190910518
step: 30, loss: 0.00028857827419415116
step: 40, loss: 0.00022133151651360095
step: 50, loss: 0.0005339114577509463
step: 60, loss: 0.00019390374654904008
step: 70, loss: 0.0004809470847249031
step: 80, loss: 0.0009695389890111983
step: 90, loss: 0.000280007574474439
step: 100, loss: 0.0011554529191926122
step: 110, loss: 0.00027141126338392496
step: 120, loss: 0.0002201939350925386
step: 130, loss: 7.848013046896085e-05
step: 140, loss: 8.59611391206272e-05
step: 150, loss: 0.00010612788901198655
step: 160, loss: 9.742394468048587e-05
step: 170, loss: 0.00021503381140064448
```

```

step: 180, loss: 0.00011249825183767825
step: 190, loss: 0.007373841479420662
/usr/local/lib/python3.7/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=0.8108108108108106, f1=0.8140161725067385,
best_f1=0.8075880758807588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014371117867995054
step: 10, loss: 0.00029329973040148616
step: 20, loss: 0.00030839178361929953
step: 30, loss: 7.280893623828888e-05
step: 40, loss: 0.0010283489245921373
step: 50, loss: 0.00014357888721860945
step: 60, loss: 0.0007205448346212506
step: 70, loss: 0.00024183224013540894
step: 80, loss: 0.00011667710350593552
step: 90, loss: 0.00021546834614127874
step: 100, loss: 0.0006186580285429955
step: 110, loss: 0.0013888961402699351
step: 120, loss: 0.00015428535698447376
step: 130, loss: 0.00028760841814801097
step: 140, loss: 0.0012558615999296308
step: 150, loss: 0.00011282821651548147
step: 160, loss: 0.00013242146815173328
step: 170, loss: 0.00014949181058909744
step: 180, loss: 0.00024197572201956064
step: 190, loss: 0.001506010303273797
/usr/local/lib/python3.7/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=0.8126649076517151, f1=0.8010471204188483,
best_f1=0.8075880758807588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016468156827613711
step: 10, loss: 0.00028425187338143587
step: 20, loss: 0.00043321819975972176
step: 30, loss: 7.590534369228408e-05
step: 40, loss: 0.00011990436178166419
step: 50, loss: 0.0003845604369416833
step: 60, loss: 0.00010772389214253053
step: 70, loss: 0.00293461955152452
step: 80, loss: 0.0002671946713235229
step: 90, loss: 0.00014044191630091518

```

```

step: 100, loss: 0.00047708849888294935
step: 110, loss: 8.873597107594833e-05
step: 120, loss: 0.00014281822950579226
step: 130, loss: 0.002310719806700945
step: 140, loss: 0.004434430971741676
step: 150, loss: 0.0004473289882298559
step: 160, loss: 7.324366742977872e-05
step: 170, loss: 0.00017870779265649617
step: 180, loss: 0.0003542936174198985
step: 190, loss: 0.00025261196424253285
/usr/local/lib/python3.7/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=0.8093994778067886, f1=0.8020833333333333,
best_f1=0.8075880758807588
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 [00:08, 239.57it/s]
load_f1 = 0.6416861826697892
real_f1 = 0.625

```

```
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 [00:17, 252.77it/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: 0.6327157020568848
step: 10, loss: 0.36531269550323486
step: 20, loss: 0.3092350959777832
step: 30, loss: 0.3933540880680084
step: 40, loss: 0.29432451725006104
step: 50, loss: 0.2873457074165344
step: 60, loss: 0.21464067697525024
step: 70, loss: 0.3646022081375122
step: 80, loss: 0.35460901260375977
step: 90, loss: 0.22150737047195435
step: 100, loss: 0.20381218194961548
step: 110, loss: 0.19730915129184723
step: 120, loss: 0.13999149203300476
step: 130, loss: 0.04296455904841423
step: 140, loss: 0.12359081953763962
step: 150, loss: 0.20664340257644653
step: 160, loss: 0.11819905787706375

```

```

step: 170, loss: 0.23480942845344543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7549999999999999, f1=0.7152941176470586,
best_f1=0.7152941176470586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04217685014009476
step: 10, loss: 0.31213298439979553
step: 20, loss: 0.04457659646868706
step: 30, loss: 0.29143768548965454
step: 40, loss: 0.05791207775473595
step: 50, loss: 0.04782060161232948
step: 60, loss: 0.11303833872079849
step: 70, loss: 0.07644245028495789
step: 80, loss: 0.05945536866784096
step: 90, loss: 0.09330316632986069
step: 100, loss: 0.16127626597881317
step: 110, loss: 0.07900188863277435
step: 120, loss: 0.06783773750066757
step: 130, loss: 0.12702985107898712
step: 140, loss: 0.3566024899482727
step: 150, loss: 0.27405163645744324
step: 160, loss: 0.16798530519008636
step: 170, loss: 0.06330536305904388
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8, f1=0.7915690866510539, best_f1=0.7915690866510539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11478451639413834
step: 10, loss: 0.060628101229667664
step: 20, loss: 0.04365130141377449
step: 30, loss: 0.1078067198395729
step: 40, loss: 0.07345747202634811
step: 50, loss: 0.0750802606344223

```

```

step: 60, loss: 0.0418398380279541
step: 70, loss: 0.14162661135196686
step: 80, loss: 0.023257702589035034
step: 90, loss: 0.08788079768419266
step: 100, loss: 0.022650035098195076
step: 110, loss: 0.06368637830018997
step: 120, loss: 0.15329012274742126
step: 130, loss: 0.04417762532830238
step: 140, loss: 0.11199158430099487
step: 150, loss: 0.017147399485111237
step: 160, loss: 0.11151716113090515
step: 170, loss: 0.03395820036530495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8283752860411899, f1=0.8026905829596412,
best_f1=0.8026905829596412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009232517331838608
step: 10, loss: 0.03290976956486702
step: 20, loss: 0.0406053364276886
step: 30, loss: 0.11154165863990784
step: 40, loss: 0.008466511964797974
step: 50, loss: 0.04638226330280304
step: 60, loss: 0.09049646556377411
step: 70, loss: 0.018088359385728836
step: 80, loss: 0.05918751284480095
step: 90, loss: 0.07879678159952164
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.13407252728939056
step: 110, loss: 0.07783501595258713
step: 120, loss: 0.010226793587207794
step: 130, loss: 0.040567271411418915
step: 140, loss: 0.05389115586876869
step: 150, loss: 0.2498641461133957
step: 160, loss: 0.032563745975494385
step: 170, loss: 0.004732317291200161
/usr/local/lib/python3.7/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=0.8167053364269141, f1=0.8256880733944955,
best_f1=0.8026905829596412

```

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

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

```
step: 0, loss: 0.01654909737408161  
step: 10, loss: 0.003144309390336275  
step: 20, loss: 0.013673839159309864  
step: 30, loss: 0.017098579555749893  
step: 40, loss: 0.04894805699586868  
step: 50, loss: 0.010930915363132954  
step: 60, loss: 0.020818503573536873  
step: 70, loss: 0.10795481503009796  
step: 80, loss: 0.04559915140271187  
step: 90, loss: 0.06893905252218246  
step: 100, loss: 0.0026597657706588507  
step: 110, loss: 0.019910987466573715  
step: 120, loss: 0.06275559961795807  
step: 130, loss: 0.0433238223195076  
step: 140, loss: 0.03510293737053871  
step: 150, loss: 0.022338496521115303  
step: 160, loss: 0.029857611283659935  
step: 170, loss: 0.020818864926695824
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not 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 5: dev_f1=0.8333333333333333, f1=0.8431372549019608,  
best_f1=0.8431372549019608
```

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

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

```
step: 0, loss: 0.001111155259422958  
step: 10, loss: 0.0017044252017512918  
step: 20, loss: 0.0016014864668250084  
step: 30, loss: 0.0005066693411208689  
step: 40, loss: 0.001751927426084876  
step: 50, loss: 0.08643089234828949  
step: 60, loss: 0.005074200686067343  
step: 70, loss: 0.05394937843084335  
step: 80, loss: 0.0015886302571743727  
step: 90, loss: 0.0008262335904873908  
step: 100, loss: 0.035829126834869385  
step: 110, loss: 0.006675995420664549  
step: 120, loss: 0.00951953325420618  
step: 130, loss: 0.0028222803957760334  
step: 140, loss: 0.05509281903505325
```

```

step: 150, loss: 0.025416726246476173
step: 160, loss: 0.05675971135497093
step: 170, loss: 0.006491989828646183
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8349056603773586, f1=0.7982456140350878,
best_f1=0.7982456140350878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005247510853223503
step: 10, loss: 0.005717999301850796
step: 20, loss: 0.0006566207739524543
step: 30, loss: 0.030207905918359756
step: 40, loss: 0.014932687394320965
step: 50, loss: 0.0006632408476434648
step: 60, loss: 0.015274839475750923
step: 70, loss: 0.12144586443901062
step: 80, loss: 0.006172560155391693
step: 90, loss: 0.00021283926616888493
step: 100, loss: 0.0012130157556384802
step: 110, loss: 0.01400814764201641
step: 120, loss: 0.0004719905264209956
step: 130, loss: 0.11132223904132843
step: 140, loss: 0.0006095209973864257
step: 150, loss: 0.00021384026331361383
step: 160, loss: 0.0003052606771234423
step: 170, loss: 0.04714532196521759
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8467153284671532, f1=0.8127853881278538,
best_f1=0.8127853881278538
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008919493295252323
step: 10, loss: 0.014038396999239922

```

```

step: 20, loss: 0.00016891484847292304
step: 30, loss: 0.017781337723135948
step: 40, loss: 0.00022511635324917734
step: 50, loss: 0.00045567721826955676
step: 60, loss: 0.027399126440286636
step: 70, loss: 0.0003316929214634001
step: 80, loss: 0.00987751130014658
step: 90, loss: 0.002611620118841529
step: 100, loss: 0.029684731736779213
step: 110, loss: 0.059606362134218216
step: 120, loss: 0.002449663821607828
step: 130, loss: 0.02429395541548729
step: 140, loss: 0.000487230223370716
step: 150, loss: 0.002481956034898758
step: 160, loss: 0.05759888142347336
step: 170, loss: 0.0031808989588171244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.85012285012285, f1=0.8321513002364066,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023330487310886383
step: 10, loss: 0.03413321077823639
step: 20, loss: 0.00038979604141786695
step: 30, loss: 0.0014638588763773441
step: 40, loss: 0.0014301768969744444
step: 50, loss: 0.00018758821533992887
step: 60, loss: 0.006234925240278244
step: 70, loss: 0.0013183718547224998
step: 80, loss: 0.0003799020196311176
step: 90, loss: 0.014661536552011967
step: 100, loss: 0.0014946568990126252
step: 110, loss: 0.011720916256308556
step: 120, loss: 0.0014202408492565155
step: 130, loss: 0.07391781359910965
step: 140, loss: 0.00013288663467392325
step: 150, loss: 0.0003039235598407686
step: 160, loss: 0.004881705157458782
step: 170, loss: 0.0032173246145248413
/usr/local/lib/python3.7/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=0.8470588235294118, f1=0.8394495412844037,
best_f1=0.8321513002364066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1353481113910675
step: 10, loss: 0.0004633025382645428
step: 20, loss: 0.06266142427921295
step: 30, loss: 0.00229503121227026
step: 40, loss: 0.00027335385675542057
step: 50, loss: 0.05096307769417763
step: 60, loss: 0.027717838063836098
step: 70, loss: 0.028549617156386375
step: 80, loss: 0.0021363121923059225
step: 90, loss: 0.010410168208181858
step: 100, loss: 0.00012452600640244782
step: 110, loss: 0.0020059356465935707
step: 120, loss: 0.00028371988446451724
step: 130, loss: 0.010742899030447006
step: 140, loss: 0.002473818603903055
step: 150, loss: 0.01143861748278141
step: 160, loss: 0.01088067889213562
step: 170, loss: 0.0035883337259292603
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8509615384615384, 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: 0.03553936257958412
step: 10, loss: 0.009462614543735981
step: 20, loss: 0.00025815152912400663
step: 30, loss: 0.00010578316869214177
step: 40, loss: 0.0029812660068273544
step: 50, loss: 0.004160432144999504
step: 60, loss: 0.0031154369935393333
step: 70, loss: 0.016423212364315987
step: 80, loss: 8.30626959213987e-05
step: 90, loss: 0.0001109883887693286
step: 100, loss: 0.00029765404178760946
step: 110, loss: 0.008874857798218727

```

```

step: 120, loss: 0.0028383515309542418
step: 130, loss: 9.914330439642072e-05
step: 140, loss: 0.004138662479817867
step: 150, loss: 0.005653894040733576
step: 160, loss: 0.0015706161502748728
step: 170, loss: 0.00023376819444820285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8530120481927711, f1=0.8356164383561643,
best_f1=0.8356164383561643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005142293404787779
step: 10, loss: 0.0007000255864113569
step: 20, loss: 0.000458532857010141
step: 30, loss: 0.0007148711592890322
step: 40, loss: 9.582657366991043e-05
step: 50, loss: 6.77548669045791e-05
step: 60, loss: 0.004948748741298914
step: 70, loss: 0.06543516367673874
step: 80, loss: 7.262474537128583e-05
step: 90, loss: 0.0008351699216291308
step: 100, loss: 0.0004627132148016244
step: 110, loss: 0.0003318128001410514
step: 120, loss: 0.0064780861139297485
step: 130, loss: 0.00795083586126566
step: 140, loss: 0.015862558037042618
step: 150, loss: 0.023327216506004333
step: 160, loss: 0.00014685335918329656
step: 170, loss: 0.00017754625878296793
/usr/local/lib/python3.7/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=0.8428571428571429, f1=0.8200455580865603,
best_f1=0.8356164383561643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0162893645465374
step: 10, loss: 0.00025952424039132893
step: 20, loss: 0.0004248158656992018
step: 30, loss: 6.375482917064801e-05

```



```

step: 40, loss: 0.0011377556947991252
step: 50, loss: 0.0001304147590417415
step: 60, loss: 0.006358208134770393
step: 70, loss: 0.00011315029405523092
step: 80, loss: 0.006604596972465515
step: 90, loss: 9.131318802246824e-05
step: 100, loss: 0.0002936789533123374
step: 110, loss: 7.992313476279378e-05
step: 120, loss: 0.046340640634298325
step: 130, loss: 0.0002740605268627405
step: 140, loss: 0.0010528047569096088
step: 150, loss: 0.010136273689568043
step: 160, loss: 0.0013825934147462249
step: 170, loss: 0.0007284996681846678
/usr/local/lib/python3.7/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=0.8402948402948404, f1=0.8345323741007193,
best_f1=0.8356164383561643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018853494839277118
step: 10, loss: 5.728998439735733e-05
step: 20, loss: 0.0008988112676888704
step: 30, loss: 0.0005692348349839449
step: 40, loss: 0.0001150903117377311
step: 50, loss: 0.00022393843391910195
step: 60, loss: 6.393033254425973e-05
step: 70, loss: 0.00043801931315101683
step: 80, loss: 0.005986290983855724
step: 90, loss: 6.0639726143563166e-05
step: 100, loss: 6.764371937606484e-05
step: 110, loss: 0.00046711089089512825
step: 120, loss: 0.030726974830031395
step: 130, loss: 0.0002571306540630758
step: 140, loss: 0.01612377166748047
step: 150, loss: 0.0004443606885615736
step: 160, loss: 6.994685099925846e-05
step: 170, loss: 0.000526642834302038
/usr/local/lib/python3.7/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=0.8428927680798006, f1=0.8345323741007193,
best_f1=0.8356164383561643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0001078166751540266
step: 10, loss: 0.002550478558987379
step: 20, loss: 0.03455658257007599
step: 30, loss: 0.0004127366410102695
step: 40, loss: 0.0002458541712258011
step: 50, loss: 0.00012415986566338688
step: 60, loss: 0.014834990724921227
step: 70, loss: 9.428390330867842e-05
step: 80, loss: 0.026513736695051193
step: 90, loss: 0.0007468058611266315
step: 100, loss: 0.00039190828101709485
step: 110, loss: 9.524391498416662e-05
step: 120, loss: 0.0018874667584896088
step: 130, loss: 0.0015817261300981045
step: 140, loss: 0.00018387612362857908
step: 150, loss: 0.0003981141489930451
step: 160, loss: 0.00035662922891788185
step: 170, loss: 7.014181755948812e-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=0.8413461538461537, f1=0.8186046511627907,
best_f1=0.8356164383561643
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:06, 285.03it/s]
load_f1 = 0.526984126984127
real_f1 = 0.5
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:17, 254.70it/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: 0.6226093769073486  
step: 10, loss: 0.6111717224121094  
step: 20, loss: 0.3302847146987915  
step: 30, loss: 0.09897974878549576  
step: 40, loss: 0.2984960079193115  
step: 50, loss: 0.08655054867267609  
step: 60, loss: 0.08739476650953293  
step: 70, loss: 0.030239932239055634
```

```

step: 80, loss: 0.0939968079328537
step: 90, loss: 0.16360080242156982
step: 100, loss: 0.008809146471321583
step: 110, loss: 0.24078230559825897
step: 120, loss: 0.007199554704129696
step: 130, loss: 0.016647515818476677
step: 140, loss: 0.001760840299539268
step: 150, loss: 0.008232886902987957
step: 160, loss: 0.010984229855239391
step: 170, loss: 0.1564292162656784
step: 180, loss: 0.05340570956468582
step: 190, loss: 0.10326531529426575
step: 200, loss: 0.060336388647556305
step: 210, loss: 0.013973123393952847
step: 220, loss: 0.016158854588866234
step: 230, loss: 0.07382582873106003
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831649831649831, f1=0.9683257918552037,
best_f1=0.9683257918552037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003918757662177086
step: 10, loss: 0.004141403362154961
step: 20, loss: 0.10170432180166245
step: 30, loss: 0.21826054155826569
step: 40, loss: 0.11569038033485413
step: 50, loss: 0.010267764329910278
step: 60, loss: 0.006646883208304644
step: 70, loss: 0.008229710161685944
step: 80, loss: 0.04305795207619667
step: 90, loss: 0.002329290611669421
step: 100, loss: 0.02019750140607357
step: 110, loss: 0.02002468705177307
step: 120, loss: 0.12923428416252136
step: 130, loss: 0.02773568220436573
step: 140, loss: 0.0026049308944493532
step: 150, loss: 0.06284692883491516
step: 160, loss: 0.03592393547296524
step: 170, loss: 0.0009677445632405579
step: 180, loss: 0.07622823119163513
step: 190, loss: 0.004258070141077042

```

```

step: 200, loss: 0.006509997881948948
step: 210, loss: 0.0017402954399585724
step: 220, loss: 0.12030035257339478
step: 230, loss: 0.008052131161093712
/usr/local/lib/python3.7/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=0.9728506787330317, f1=0.971815107102593,
best_f1=0.9683257918552037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03517082333564758
step: 10, loss: 0.003559060860425234
step: 20, loss: 0.036658212542533875
step: 30, loss: 0.06273113936185837
step: 40, loss: 0.05225203558802605
step: 50, loss: 0.005856599658727646
step: 60, loss: 0.003101054346188903
step: 70, loss: 0.002177830785512924
step: 80, loss: 0.0015706628328189254
step: 90, loss: 0.08545278012752533
step: 100, loss: 0.007047642488032579
step: 110, loss: 0.0023399183992296457
step: 120, loss: 0.030559374019503593
step: 130, loss: 0.005893499590456486
step: 140, loss: 0.00496146734803915
step: 150, loss: 0.06007017195224762
step: 160, loss: 0.036019884049892426
step: 170, loss: 0.008945081382989883
step: 180, loss: 0.003026788355782628
step: 190, loss: 0.014355885796248913
step: 200, loss: 0.003319921437650919
step: 210, loss: 0.01167785283178091
step: 220, loss: 0.0009971152758225799
step: 230, loss: 0.0016977089690044522
/usr/local/lib/python3.7/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=0.9799554565701558, f1=0.9787709497206705,
best_f1=0.9683257918552037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00034445725032128394
step: 10, loss: 0.0054124402813613415
step: 20, loss: 0.0008945303270593286
step: 30, loss: 0.0006621414213441312

```

```

step: 40, loss: 0.002516008447855711
step: 50, loss: 0.0004255493695382029
step: 60, loss: 0.001522220321930945
step: 70, loss: 0.0007666131714358926
step: 80, loss: 0.004051056690514088
step: 90, loss: 0.001176622579805553
step: 100, loss: 0.0017975844675675035
step: 110, loss: 0.0008157754200510681
step: 120, loss: 0.009556268341839314
step: 130, loss: 0.003319663228467107
step: 140, loss: 0.0007262454018928111
step: 150, loss: 0.1408979296684265
step: 160, loss: 0.010678382590413094
step: 170, loss: 0.004957861732691526
step: 180, loss: 0.0006460549775511026
step: 190, loss: 0.004561914596706629
step: 200, loss: 0.002936332719400525
step: 210, loss: 0.016454674303531647
step: 220, loss: 0.0003797622921410948
step: 230, loss: 0.008388658054172993
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9844097995545658, 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: 0.0017497276421636343
step: 10, loss: 0.0005535512464120984
step: 20, loss: 0.0003072464023716748
step: 30, loss: 0.0002769680868368596
step: 40, loss: 0.002170293591916561
step: 50, loss: 0.00022808393987361342
step: 60, loss: 0.12868434190750122
step: 70, loss: 0.001632018480449915
step: 80, loss: 0.0005628517246805131
step: 90, loss: 0.002088701818138361
step: 100, loss: 0.00024044522433541715
step: 110, loss: 0.004876713268458843
step: 120, loss: 0.0001552915055071935
step: 130, loss: 0.0007898146868683398
step: 140, loss: 0.004581528715789318
step: 150, loss: 0.0005102395662106574

```

```

step: 160, loss: 0.003388448152691126
step: 170, loss: 0.0021431115455925465
step: 180, loss: 0.0006504315533675253
step: 190, loss: 0.08681057393550873
step: 200, loss: 0.0027942457236349583
step: 210, loss: 0.002119645709171891
step: 220, loss: 0.0013656606897711754
step: 230, loss: 0.0003922964970115572
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9732739420935412,
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: 0.0019221603870391846
step: 10, loss: 0.0007694970117881894
step: 20, loss: 0.0011127879843115807
step: 30, loss: 0.00018171474221162498
step: 40, loss: 0.00018657141481526196
step: 50, loss: 0.0004012440040241927
step: 60, loss: 0.002015538979321718
step: 70, loss: 0.005032976157963276
step: 80, loss: 0.001870288047939539
step: 90, loss: 0.0004581291868817061
step: 100, loss: 0.06955563277006149
step: 110, loss: 0.0012546718353405595
step: 120, loss: 0.001523603918030858
step: 130, loss: 0.0012464680476114154
step: 140, loss: 0.0003770476614590734
step: 150, loss: 0.010950231924653053
step: 160, loss: 0.000300076964776963
step: 170, loss: 0.0008153424714691937
step: 180, loss: 0.011999289505183697
step: 190, loss: 0.003424376714974642
step: 200, loss: 0.007947418838739395
step: 210, loss: 0.0005666849901899695
step: 220, loss: 0.00028915569419041276
step: 230, loss: 0.07716761529445648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854748603351955, f1=0.9776286353467561,

```



```

best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005635290290229023
step: 10, loss: 0.0001772209070622921
step: 20, loss: 0.0007519427454099059
step: 30, loss: 0.0009541361359879375
step: 40, loss: 0.0015013439115136862
step: 50, loss: 0.0004038449842482805
step: 60, loss: 0.0030287159606814384
step: 70, loss: 0.02347848378121853
step: 80, loss: 0.0021258259657770395
step: 90, loss: 0.0001387364900438115
step: 100, loss: 0.00016976401093415916
step: 110, loss: 0.0002002044057007879
step: 120, loss: 7.781679596519098e-05
step: 130, loss: 0.0006599330808967352
step: 140, loss: 0.005186321213841438
step: 150, loss: 0.004121905192732811
step: 160, loss: 0.06935902684926987
step: 170, loss: 0.0010325150797143579
step: 180, loss: 0.002027973998337984
step: 190, loss: 0.0005494979559443891
step: 200, loss: 0.04933497682213783
step: 210, loss: 9.637260518502444e-05
step: 220, loss: 0.000306903850287199
step: 230, loss: 0.00017143625882454216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9855072463768116, f1=0.9745293466223698,
best_f1=0.9745293466223698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017180350550916046
step: 10, loss: 0.00040088690002448857
step: 20, loss: 0.0003373404033482075
step: 30, loss: 0.022772736847400665
step: 40, loss: 0.0004777696740347892
step: 50, loss: 0.000580832886043936
step: 60, loss: 0.00014297547750175
step: 70, loss: 0.00016894213331397623

```

```

step: 80, loss: 0.0001435179146938026
step: 90, loss: 9.290342131862417e-05
step: 100, loss: 0.0001528501888969913
step: 110, loss: 0.0002750819840002805
step: 120, loss: 0.004493183922022581
step: 130, loss: 0.00014084286522120237
step: 140, loss: 8.002013055374846e-05
step: 150, loss: 0.007078732829540968
step: 160, loss: 0.0003647159901447594
step: 170, loss: 0.0003680457011796534
step: 180, loss: 0.0005542451981455088
step: 190, loss: 0.0011036940850317478
step: 200, loss: 0.00215321546420455
step: 210, loss: 0.00044298148714005947
step: 220, loss: 0.0004380153550300747
step: 230, loss: 0.0014247811632230878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865470852017937, f1=0.9797752808988766,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.479373016394675e-05
step: 10, loss: 0.001761177321895957
step: 20, loss: 0.0011414560722187161
step: 30, loss: 0.00014512201596517116
step: 40, loss: 0.009573245421051979
step: 50, loss: 7.220856787171215e-05
step: 60, loss: 0.0001772063405951485
step: 70, loss: 0.00045466338633559644
step: 80, loss: 0.00011021206591976807
step: 90, loss: 0.00228897831402719
step: 100, loss: 0.00014852784806862473
step: 110, loss: 6.40752405161038e-05
step: 120, loss: 6.162296631373465e-05
step: 130, loss: 0.00010799283336382359
step: 140, loss: 6.158422911539674e-05
step: 150, loss: 0.00020752695854753256
step: 160, loss: 6.919638690305874e-05
step: 170, loss: 0.0001263085869140923
step: 180, loss: 0.00026154849911108613
step: 190, loss: 0.0001243221340700984

```

```

step: 200, loss: 0.00013440073234960437
step: 210, loss: 0.00010273139923810959
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 220, loss: 6.831907376181334e-05
step: 230, loss: 7.959338836371899e-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 9: dev_f1=0.9876543209876544, f1=0.9797752808988766,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.65495894686319e-05
step: 10, loss: 5.618697468889877e-05
step: 20, loss: 0.00014314927102532238
step: 30, loss: 0.0006871780497021973
step: 40, loss: 7.80061018303968e-05
step: 50, loss: 0.000352442788425833
step: 60, loss: 8.171743684215471e-05
step: 70, loss: 0.00018762746185529977
step: 80, loss: 6.954406126169488e-05
step: 90, loss: 0.00011260404426138848
step: 100, loss: 5.708239405066706e-05
step: 110, loss: 6.763399869669229e-05
step: 120, loss: 0.00011832798918476328
step: 130, loss: 0.07361533492803574
step: 140, loss: 0.04671861231327057
step: 150, loss: 0.011776918545365334
step: 160, loss: 0.000138312199851498
step: 170, loss: 0.00029476359486579895
step: 180, loss: 0.0011183419264853
step: 190, loss: 0.011742661707103252
step: 200, loss: 5.2249772124923766e-05
step: 210, loss: 7.104055111994967e-05
step: 220, loss: 5.731330384151079e-05
step: 230, loss: 0.00019822161993943155
/usr/local/lib/python3.7/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=0.9865771812080537, f1=0.9777777777777777,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003193129086866975
step: 10, loss: 0.00011523177090566605
step: 20, loss: 0.0009606968378648162
step: 30, loss: 0.000382504949811846
step: 40, loss: 0.0008144428138621151
step: 50, loss: 7.187854498624802e-05
step: 60, loss: 0.000570887525100261
step: 70, loss: 8.371525473194197e-05
step: 80, loss: 6.945180211914703e-05
step: 90, loss: 5.27303091075737e-05
step: 100, loss: 6.8582572566811e-05
step: 110, loss: 0.02594638615846634
step: 120, loss: 4.2890958866337314e-05
step: 130, loss: 4.2161896999459714e-05
step: 140, loss: 3.589900370570831e-05
step: 150, loss: 0.041673555970191956
step: 160, loss: 3.976185689680278e-05
step: 170, loss: 0.02156383916735649
step: 180, loss: 0.00010940910578938201
step: 190, loss: 0.00017945116269402206
step: 200, loss: 0.001970843877643347
step: 210, loss: 0.0001347535871900618
step: 220, loss: 7.39816969144158e-05
step: 230, loss: 5.268950553727336e-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=0.9865771812080537, f1=0.9777777777777777,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.560962356161326e-05
step: 10, loss: 4.6863719035172835e-05
step: 20, loss: 4.353091571829282e-05
step: 30, loss: 0.000389053369872272
step: 40, loss: 8.571577200200409e-05
step: 50, loss: 0.0005780279752798378
step: 60, loss: 0.00011211131641175598
step: 70, loss: 5.490149123943411e-05
step: 80, loss: 0.0001264384773094207
step: 90, loss: 3.630861101555638e-05
step: 100, loss: 3.113813363597728e-05
step: 110, loss: 4.0986800740938634e-05
step: 120, loss: 0.00011169906065333635
step: 130, loss: 5.35260587639641e-05
step: 140, loss: 6.222701631486416e-05

```

```

step: 150, loss: 9.577913442626595e-05
step: 160, loss: 7.008135435171425e-05
step: 170, loss: 9.005281026475132e-05
step: 180, loss: 0.0006014920654706657
step: 190, loss: 5.225511995377019e-05
step: 200, loss: 2.4898979972931556e-05
step: 210, loss: 4.5766366383759305e-05
step: 220, loss: 3.0274242817540653e-05
step: 230, loss: 0.010973779484629631
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9798206278026906,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017061375547200441
step: 10, loss: 8.316356979776174e-05
step: 20, loss: 6.665520049864426e-05
step: 30, loss: 8.070833428064361e-05
step: 40, loss: 0.00017061502148862928
step: 50, loss: 7.035465387161821e-05
step: 60, loss: 4.70212908112444e-05
step: 70, loss: 3.243429819121957e-05
step: 80, loss: 6.154508446343243e-05
step: 90, loss: 0.006275487132370472
step: 100, loss: 0.0001040477873175405
step: 110, loss: 0.019353916868567467
step: 120, loss: 0.026941752061247826
step: 130, loss: 6.234062311705202e-05
step: 140, loss: 4.534537220024504e-05
step: 150, loss: 5.173291719984263e-05
step: 160, loss: 0.0015315302880480886
step: 170, loss: 9.168030373984948e-05
step: 180, loss: 4.022005668957718e-05
step: 190, loss: 4.024159352411516e-05
step: 200, loss: 0.00036964999162591994
step: 210, loss: 2.2351130610331893e-05
step: 220, loss: 4.1842940845526755e-05
step: 230, loss: 4.484094097279012e-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=0.975609756097561,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.5973614760441706e-05
step: 10, loss: 7.391626422759145e-05
step: 20, loss: 5.3246425522957e-05
step: 30, loss: 6.420638965209946e-05
step: 40, loss: 0.00894906185567379
step: 50, loss: 0.00011101910786237568
step: 60, loss: 5.0547347200335935e-05
step: 70, loss: 4.4533266191137955e-05
step: 80, loss: 0.00010594517516437918
step: 90, loss: 4.871951023233123e-05
step: 100, loss: 4.730289219878614e-05
step: 110, loss: 0.00011715287837432697
step: 120, loss: 2.6679557777242735e-05
step: 130, loss: 4.346052082837559e-05
step: 140, loss: 4.467507096705958e-05
step: 150, loss: 3.2188927434617653e-05
step: 160, loss: 0.0017280426109209657
step: 170, loss: 2.0678600776591338e-05
step: 180, loss: 4.909516792395152e-05
step: 190, loss: 0.00023544422583654523
step: 200, loss: 4.515661930781789e-05
step: 210, loss: 0.0001045737590175122
step: 220, loss: 7.52556006773375e-05
step: 230, loss: 0.0004050623392686248
/usr/local/lib/python3.7/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=0.9843749999999999, f1=0.9777777777777777,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.5320050554000773e-05
step: 10, loss: 1.8883107259171084e-05
step: 20, loss: 6.160042539704591e-05
step: 30, loss: 0.00013367991778068244
step: 40, loss: 0.00022349607024807483
step: 50, loss: 0.03795769438147545
step: 60, loss: 4.5211265387479216e-05
step: 70, loss: 0.00020194465469103307
step: 80, loss: 3.810325506492518e-05
step: 90, loss: 7.490545249311253e-05
step: 100, loss: 5.523057916434482e-05
step: 110, loss: 2.5748464395292103e-05
step: 120, loss: 5.6814809795469046e-05
step: 130, loss: 5.646925637847744e-05
step: 140, loss: 3.1969342671800405e-05

```

```

step: 150, loss: 7.401720358757302e-05
step: 160, loss: 3.1297335226554424e-05
step: 170, loss: 2.8914331778651103e-05
step: 180, loss: 7.743819878669456e-05
step: 190, loss: 4.454022928257473e-05
step: 200, loss: 8.611290832050145e-05
step: 210, loss: 0.00012267817510291934
step: 220, loss: 0.00034180161310359836
step: 230, loss: 4.208893733448349e-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=0.9854096520763187, f1=0.9786276715410572,
best_f1=0.9797752808988766
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 [00:11, 207.53it/s]
load_f1 = 0.9876543209876544
real_f1 = 0.9865470852017937
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 [00:17, 251.12it/s]
```

3.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6280346512794495
step: 10, loss: 0.5080435276031494
step: 20, loss: 0.4767339825630188
step: 30, loss: 0.10221357643604279
step: 40, loss: 0.1682855486869812
step: 50, loss: 0.13979434967041016
step: 60, loss: 0.04335460439324379
step: 70, loss: 0.12948046624660492
step: 80, loss: 0.03186746686697006
step: 90, loss: 0.10679835826158524
step: 100, loss: 0.016919391229748726
step: 110, loss: 0.03343882039189339
step: 120, loss: 0.11478377133607864
step: 130, loss: 0.08298378437757492
step: 140, loss: 0.13990698754787445
step: 150, loss: 0.019931750372052193
step: 160, loss: 0.012511574663221836
step: 170, loss: 0.23464079201221466
step: 180, loss: 0.0870489552617073
step: 190, loss: 0.006113993003964424

```

```

step: 200, loss: 0.18741755187511444
step: 210, loss: 0.07614054530858994
step: 220, loss: 0.27904918789863586
step: 230, loss: 0.14220237731933594
step: 240, loss: 0.03312020003795624
step: 250, loss: 0.018048835918307304
step: 260, loss: 0.08776874095201492
step: 270, loss: 0.02424396201968193
step: 280, loss: 0.06517990678548813
step: 290, loss: 0.027090637013316154
step: 300, loss: 0.019029341638088226
step: 310, loss: 0.09741830080747604
step: 320, loss: 0.10177161544561386
step: 330, loss: 0.017824506387114525
step: 340, loss: 0.027645939961075783
step: 350, loss: 0.009257986210286617
step: 360, loss: 0.017633752897381783
step: 370, loss: 0.12040811032056808
step: 380, loss: 0.015983525663614273
step: 390, loss: 0.08840951323509216
step: 400, loss: 0.3814774453639984
step: 410, loss: 0.05105984956026077
step: 420, loss: 0.15338727831840515
step: 430, loss: 0.1771610677242279
step: 440, loss: 0.049275953322649
step: 450, loss: 0.0027759065851569176
step: 460, loss: 0.004361357074230909
step: 470, loss: 0.1296318769454956
step: 480, loss: 0.031195366755127907
step: 490, loss: 0.06278689950704575
step: 500, loss: 0.0773644894361496
step: 510, loss: 0.08139947056770325
step: 520, loss: 0.0337681844830513
step: 530, loss: 0.00667008850723505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9391385767790261, f1=0.9400749063670412,
best_f1=0.9400749063670412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12527018785476685
step: 10, loss: 0.03922409564256668

```

step: 20, loss: 0.011397373862564564
step: 30, loss: 0.004772004671394825
step: 40, loss: 0.2355617731809616
step: 50, loss: 0.22593820095062256
step: 60, loss: 0.022869667038321495
step: 70, loss: 0.0369827039539814
step: 80, loss: 0.07708856463432312
step: 90, loss: 0.009801337495446205
step: 100, loss: 0.014631686732172966
step: 110, loss: 0.0474032461643219
step: 120, loss: 0.10651502013206482
step: 130, loss: 0.058389998972415924
step: 140, loss: 0.023253118619322777
step: 150, loss: 0.0914546400308609
step: 160, loss: 0.0024526866618543863
step: 170, loss: 0.004643731750547886
step: 180, loss: 0.08494371175765991
step: 190, loss: 0.06106261536478996
step: 200, loss: 0.007724866736680269
step: 210, loss: 0.061490535736083984
step: 220, loss: 0.09076842665672302
step: 230, loss: 0.004530081059783697
step: 240, loss: 0.051024630665779114
step: 250, loss: 0.013305222615599632
step: 260, loss: 0.0023357588797807693
step: 270, loss: 0.19429545104503632
step: 280, loss: 0.11500705033540726
step: 290, loss: 0.0342836007475853
step: 300, loss: 0.19670511782169342
step: 310, loss: 0.012050459161400795
step: 320, loss: 0.1380360722541809
step: 330, loss: 0.03543896973133087
step: 340, loss: 0.013954417780041695
step: 350, loss: 0.002897006692364812
step: 360, loss: 0.11570046097040176
step: 370, loss: 0.05194542557001114
step: 380, loss: 0.03178810700774193
step: 390, loss: 0.0651790127158165
step: 400, loss: 0.013699742034077644
step: 410, loss: 0.00940902903676033
step: 420, loss: 0.030888814479112625
step: 430, loss: 0.006113944575190544
step: 440, loss: 0.2755654752254486
step: 450, loss: 0.01370586920529604
step: 460, loss: 0.019149499014019966
step: 470, loss: 0.11234951764345169
step: 480, loss: 0.2234613448381424
step: 490, loss: 0.10112457722425461

```

step: 500, loss: 0.22987864911556244
step: 510, loss: 0.013003945350646973
step: 520, loss: 0.03992634639143944
step: 530, loss: 0.008437315933406353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9466911764705882, f1=0.9391705069124424,
best_f1=0.9391705069124424
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016155453398823738
step: 10, loss: 0.04351814463734627
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.09378373622894287
step: 30, loss: 0.09705734997987747
step: 40, loss: 0.0738321915268898
step: 50, loss: 0.015054692514240742
step: 60, loss: 0.031009193509817123
step: 70, loss: 0.005456792190670967
step: 80, loss: 0.0006548622041009367
step: 90, loss: 0.006658511236310005
step: 100, loss: 0.09404424577951431
step: 110, loss: 0.002488359808921814
step: 120, loss: 0.01247132383286953
step: 130, loss: 0.0022088014520704746
step: 140, loss: 0.07580810785293579
step: 150, loss: 0.023361176252365112
step: 160, loss: 0.031749606132507324
step: 170, loss: 0.09099229425191879
step: 180, loss: 0.14926978945732117
step: 190, loss: 0.00547253480181098
step: 200, loss: 0.04397162050008774
step: 210, loss: 0.04198308289051056
step: 220, loss: 0.064799003303051
step: 230, loss: 0.10562896728515625
step: 240, loss: 0.0018995744176208973
step: 250, loss: 0.028092138469219208
step: 260, loss: 0.009384648874402046
step: 270, loss: 0.0014643609756603837
step: 280, loss: 0.20751695334911346
step: 290, loss: 0.004480505827814341
step: 300, loss: 0.11472269147634506

```

```

step: 310, loss: 0.001697855768725276
step: 320, loss: 0.0061485059559345245
step: 330, loss: 0.0010941139189526439
step: 340, loss: 0.004048462491482496
step: 350, loss: 0.001083084149286151
step: 360, loss: 0.030881524085998535
step: 370, loss: 0.024127071723341942
step: 380, loss: 0.018292343243956566
step: 390, loss: 0.009055704809725285
step: 400, loss: 0.01243430096656084
step: 410, loss: 0.021192215383052826
step: 420, loss: 0.22945134341716766
step: 430, loss: 0.011649714782834053
step: 440, loss: 0.00445073377341032
step: 450, loss: 0.10716671496629715
step: 460, loss: 0.04925533011555672
step: 470, loss: 0.08546806871891022
step: 480, loss: 0.0025175034534186125
step: 490, loss: 0.006115547381341457
step: 500, loss: 0.030329585075378418
step: 510, loss: 0.001969777513295412
step: 520, loss: 0.022446461021900177
step: 530, loss: 0.08862382918596268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9499766245909304, f1=0.9419475655430711,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010371200740337372
step: 10, loss: 0.017721762880682945
step: 20, loss: 0.0028650679159909487
step: 30, loss: 0.003943332005292177
step: 40, loss: 0.0008575418032705784
step: 50, loss: 0.011982194148004055
step: 60, loss: 0.004004329442977905
step: 70, loss: 0.004628316964954138
step: 80, loss: 0.0036442356649786234
step: 90, loss: 0.01153544895350933
step: 100, loss: 0.0034695256035774946
step: 110, loss: 0.028980476781725883
step: 120, loss: 0.00028580761863850057

```

```

step: 130, loss: 0.0026623632293194532
step: 140, loss: 0.0010191481560468674
step: 150, loss: 0.017631690949201584
step: 160, loss: 0.0006468482897616923
step: 170, loss: 0.005648661870509386
step: 180, loss: 0.00028096180176362395
step: 190, loss: 0.002476150169968605
step: 200, loss: 0.007251759059727192
step: 210, loss: 0.01290825754404068
step: 220, loss: 0.007271878886967897
step: 230, loss: 0.08596759289503098
step: 240, loss: 0.0018167621456086636
step: 250, loss: 0.005991328973323107
step: 260, loss: 0.007191129494458437
step: 270, loss: 0.007670114748179913
step: 280, loss: 0.008749718777835369
step: 290, loss: 0.016469206660985947
step: 300, loss: 0.00016168420552276075
step: 310, loss: 0.0028687308076769114
step: 320, loss: 0.005699171219021082
step: 330, loss: 0.004259136039763689
step: 340, loss: 0.0035061410162597895
step: 350, loss: 0.07238981872797012
step: 360, loss: 0.014724202454090118
step: 370, loss: 0.006540300790220499
step: 380, loss: 0.0018711372977122664
step: 390, loss: 0.0006298920488916337
step: 400, loss: 0.001843899255618453
step: 410, loss: 0.0025042316410690546
step: 420, loss: 0.0007256873068399727
step: 430, loss: 0.12594948709011078
step: 440, loss: 0.0008331143762916327
step: 450, loss: 0.0022214192431420088
step: 460, loss: 0.0012169405817985535
step: 470, loss: 0.006427224259823561
step: 480, loss: 0.1289213001728058
step: 490, loss: 0.002914735581725836
step: 500, loss: 0.0061530908569693565
step: 510, loss: 0.006434346549212933
step: 520, loss: 0.11924498528242111
step: 530, loss: 0.02293069288134575
/usr/local/lib/python3.7/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=0.9447236180904524, f1=0.9387568555758683,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010676013305783272
step: 10, loss: 0.04975922778248787
step: 20, loss: 0.0017258374718949199
step: 30, loss: 0.09314175695180893
step: 40, loss: 0.0020306736696511507
step: 50, loss: 0.0007500086212530732
step: 60, loss: 0.021380139514803886
step: 70, loss: 0.008671828545629978
step: 80, loss: 0.001259667449630797
step: 90, loss: 0.010233437642455101
step: 100, loss: 0.0012284740805625916
step: 110, loss: 0.0006453974056057632
step: 120, loss: 0.0006738879019394517
step: 130, loss: 0.0003029807412531227
step: 140, loss: 0.0008810767903923988
step: 150, loss: 0.012211533263325691
step: 160, loss: 0.0002879879903048277
step: 170, loss: 0.019793590530753136
step: 180, loss: 0.0005708449753001332
step: 190, loss: 0.12578007578849792
step: 200, loss: 0.0003169348929077387
step: 210, loss: 0.006339533720165491
step: 220, loss: 0.05402867868542671
step: 230, loss: 0.0051100729033350945
step: 240, loss: 0.015465094707906246
step: 250, loss: 0.0002741449570748955
step: 260, loss: 0.0016844428610056639
step: 270, loss: 0.00020436482736840844
step: 280, loss: 0.025438038632273674
step: 290, loss: 0.16847431659698486
step: 300, loss: 0.028160857036709785
step: 310, loss: 0.009945044293999672
step: 320, loss: 0.2647175192832947
step: 330, loss: 0.014627550728619099
step: 340, loss: 0.005231060553342104
step: 350, loss: 0.0031133536249399185
step: 360, loss: 0.007342669181525707
step: 370, loss: 0.009347166866064072
step: 380, loss: 0.035053230822086334
step: 390, loss: 0.00026493941550143063
step: 400, loss: 0.021366607397794724
step: 410, loss: 0.0021633438300341368
step: 420, loss: 0.009041476994752884
step: 430, loss: 0.005760315805673599
step: 440, loss: 0.04973602294921875
step: 450, loss: 0.02596130035817623
step: 460, loss: 0.0018385595176368952
```

```

step: 470, loss: 0.0440450981259346
step: 480, loss: 0.02835376188158989
step: 490, loss: 0.0006461523589678109
step: 500, loss: 0.007714597973972559
step: 510, loss: 0.010935181751847267
step: 520, loss: 0.0009746726136654615
step: 530, loss: 0.0004474734596442431
/usr/local/lib/python3.7/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=0.9426987060998151, f1=0.9339534883720929,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015535516431555152
step: 10, loss: 0.00022776948753744364
step: 20, loss: 0.14249607920646667
step: 30, loss: 0.019931873306632042
step: 40, loss: 0.0005615564878098667
step: 50, loss: 0.028529897332191467
step: 60, loss: 0.0001882223441498354
step: 70, loss: 0.0012755121570080519
step: 80, loss: 0.019329212605953217
step: 90, loss: 0.0011304551735520363
step: 100, loss: 0.005589937325567007
step: 110, loss: 0.013802519999444485
step: 120, loss: 0.00016147849964909256
step: 130, loss: 0.00014737604942638427
step: 140, loss: 0.0006279220106080174
step: 150, loss: 0.00022623575932811946
step: 160, loss: 0.0034457624424248934
step: 170, loss: 0.0012874294770881534
step: 180, loss: 0.007794308941811323
step: 190, loss: 0.0013314177049323916
step: 200, loss: 0.00025295940577052534
step: 210, loss: 0.0021303510293364525
step: 220, loss: 0.0012470697984099388
step: 230, loss: 0.0005594568210653961
step: 240, loss: 0.048052627593278885
step: 250, loss: 0.004219549708068371
step: 260, loss: 6.083916377974674e-05
step: 270, loss: 0.04116717353463173
step: 280, loss: 0.06528202444314957
step: 290, loss: 0.0021429643966257572
step: 300, loss: 0.0006304956623353064
step: 310, loss: 0.0006490199593827128
step: 320, loss: 0.006358710583299398

```



```

step: 330, loss: 0.0009275844204239547
step: 340, loss: 0.03345486894249916
step: 350, loss: 0.0016532051376998425
step: 360, loss: 0.0409683957695961
step: 370, loss: 0.000964469974860549
step: 380, loss: 7.433573046000674e-05
step: 390, loss: 0.023713070899248123
step: 400, loss: 0.00012204973609186709
step: 410, loss: 0.0038132129702717066
step: 420, loss: 0.0019361842423677444
step: 430, loss: 0.00040957663441076875
step: 440, loss: 8.927891758503392e-05
step: 450, loss: 0.0006365547887980938
step: 460, loss: 0.00011122469004476443
step: 470, loss: 0.00010060749627882615
step: 480, loss: 0.004700148943811655
step: 490, loss: 0.007781929802149534
step: 500, loss: 0.07921892404556274
step: 510, loss: 0.15333886444568634
step: 520, loss: 0.002796437358483672
step: 530, loss: 0.025593852624297142
/usr/local/lib/python3.7/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=0.9461077844311377, f1=0.9412304866850323,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033036453532986343
step: 10, loss: 0.0011589840287342668
step: 20, loss: 0.09712770581245422
step: 30, loss: 0.0009128973470069468
step: 40, loss: 0.0008353497250936925
step: 50, loss: 0.0011873902985826135
step: 60, loss: 0.0007858816534280777
step: 70, loss: 0.00042541028233245015
step: 80, loss: 0.0012390892952680588
step: 90, loss: 0.00010669062612578273
step: 100, loss: 6.311912147793919e-05
step: 110, loss: 0.0004822045157197863
step: 120, loss: 0.0004290096985641867
step: 130, loss: 0.011659149080514908
step: 140, loss: 0.00013992299500387162
step: 150, loss: 0.0014377613551914692
step: 160, loss: 0.00020990490156691521
step: 170, loss: 0.000495314656291157
step: 180, loss: 0.00013677238894160837

```

```

step: 190, loss: 0.003447805531322956
step: 200, loss: 0.00033895642263814807
step: 210, loss: 0.0018141788896173239
step: 220, loss: 0.0002758137707132846
step: 230, loss: 0.023446517065167427
step: 240, loss: 0.004626867361366749
step: 250, loss: 0.00020746528753079474
step: 260, loss: 0.00014790068962611258
step: 270, loss: 8.104267180897295e-05
step: 280, loss: 0.00016132471500895917
step: 290, loss: 8.390080620301887e-05
step: 300, loss: 0.00030086966580711305
step: 310, loss: 9.78228563326411e-05
step: 320, loss: 0.00010717572149587795
step: 330, loss: 0.0005590622895397246
step: 340, loss: 0.034846000373363495
step: 350, loss: 0.0007105288095772266
step: 360, loss: 0.00016409750969614834
step: 370, loss: 8.387128764297813e-05
step: 380, loss: 0.00038515840424224734
step: 390, loss: 0.001228943932801485
step: 400, loss: 0.020487532019615173
step: 410, loss: 0.005969849415123463
step: 420, loss: 0.001236379612237215
step: 430, loss: 5.5292577599175274e-05
step: 440, loss: 0.0002998042036779225
step: 450, loss: 0.00044413769501261413
step: 460, loss: 0.0002925279550254345
step: 470, loss: 0.012492499314248562
step: 480, loss: 0.1374579668045044
step: 490, loss: 0.0001537664793431759
step: 500, loss: 0.04005097597837448
step: 510, loss: 0.0010028351098299026
step: 520, loss: 0.0032295440323650837
step: 530, loss: 0.0013346979394555092
/usr/local/lib/python3.7/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=0.9466357308584686, f1=0.9388136384866884,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001395749277435243
step: 10, loss: 0.009434628300368786
step: 20, loss: 6.400668644346297e-05
step: 30, loss: 0.0002820705994963646
step: 40, loss: 7.32888001948595e-05

```

step: 50, loss: 0.0024353244807571173
step: 60, loss: 0.00044493956374935806
step: 70, loss: 6.240810034796596e-05
step: 80, loss: 0.00017298611055593938
step: 90, loss: 0.0006912904791533947
step: 100, loss: 5.3043659136164933e-05
step: 110, loss: 9.364241122966632e-05
step: 120, loss: 0.0004340510058682412
step: 130, loss: 0.00021786104480270296
step: 140, loss: 0.004251738078892231
step: 150, loss: 5.4790176363894716e-05
step: 160, loss: 0.0034673039335757494
step: 170, loss: 0.00024564354680478573
step: 180, loss: 6.349005707306787e-05
step: 190, loss: 0.0032163450960069895
step: 200, loss: 0.0012953337281942368
step: 210, loss: 0.001022583106532693
step: 220, loss: 0.0011626554187387228
step: 230, loss: 0.00011820114741567522
step: 240, loss: 0.0005387227865867317
step: 250, loss: 0.012234849855303764
step: 260, loss: 7.996216299943626e-05
step: 270, loss: 0.15174871683120728
step: 280, loss: 0.025311890989542007
step: 290, loss: 0.0001283908059122041
step: 300, loss: 0.0004128973523620516
step: 310, loss: 0.0003143202338833362
step: 320, loss: 0.011571800336241722
step: 330, loss: 0.004695611540228128
step: 340, loss: 0.011355265974998474
step: 350, loss: 6.061967360437848e-05
step: 360, loss: 0.0009938287548720837
step: 370, loss: 0.00012699846411123872
step: 380, loss: 0.006605100352317095
step: 390, loss: 0.07340431213378906
step: 400, loss: 9.338120435131714e-05
step: 410, loss: 7.643405115231872e-05
step: 420, loss: 0.0051482426933944225
step: 430, loss: 0.03511204198002815
step: 440, loss: 0.0019572600722312927
step: 450, loss: 0.012943601235747337
step: 460, loss: 0.000151755302795209
step: 470, loss: 0.0001235200179507956
step: 480, loss: 4.406287189340219e-05
step: 490, loss: 0.003211965085938573
step: 500, loss: 0.00022019089374225587
step: 510, loss: 0.00013913474685978144
step: 520, loss: 0.0017732378328219056

```
step: 530, loss: 0.011088969185948372
/usr/local/lib/python3.7/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=0.941447671738128, f1=0.9367205542725173,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000923911458812654
step: 10, loss: 4.3787589675048366e-05
step: 20, loss: 0.005126680713146925
step: 30, loss: 0.00028005545027554035
step: 40, loss: 0.00025970101705752313
step: 50, loss: 9.399885311722755e-05
step: 60, loss: 4.3535408622119576e-05
step: 70, loss: 0.03828491270542145
step: 80, loss: 0.014676116406917572
step: 90, loss: 0.0026367446407675743
step: 100, loss: 0.00011979672854067758
step: 110, loss: 4.012596036773175e-05
step: 120, loss: 0.00024637739988975227
step: 130, loss: 0.020300675183534622
step: 140, loss: 0.0012544519267976284
step: 150, loss: 0.007573454640805721
step: 160, loss: 0.0004181982949376106
step: 170, loss: 0.03657718002796173
step: 180, loss: 3.835218012682162e-05
step: 190, loss: 0.0007552813040092587
step: 200, loss: 8.880509267328307e-05
step: 210, loss: 0.000665382482111454
step: 220, loss: 0.00018237883341498673
step: 230, loss: 0.00019703025463968515
step: 240, loss: 3.813791408902034e-05
step: 250, loss: 0.0001317090936936438
step: 260, loss: 0.18039149045944214
step: 270, loss: 8.280898327939212e-05
step: 280, loss: 0.0015501425368711352
step: 290, loss: 0.026276279240846634
step: 300, loss: 4.283848829800263e-05
step: 310, loss: 0.02429427020251751
step: 320, loss: 0.0006402574945241213
step: 330, loss: 0.002468144055455923
step: 340, loss: 0.007953358814120293
step: 350, loss: 0.05108487233519554
step: 360, loss: 3.271247260272503e-05
step: 370, loss: 0.00034326285822317004
step: 380, loss: 5.489377144840546e-05
```

```

step: 390, loss: 3.4531261917436495e-05
step: 400, loss: 0.13606992363929749
step: 410, loss: 5.182841050554998e-05
step: 420, loss: 3.4628581488505006e-05
step: 430, loss: 7.15188798494637e-05
step: 440, loss: 0.004588825162500143
step: 450, loss: 9.005558240460232e-05
step: 460, loss: 8.835206244839355e-05
step: 470, loss: 1.9780793081736192e-05
step: 480, loss: 4.10129468946252e-05
step: 490, loss: 0.006989636924117804
step: 500, loss: 0.0007553467876277864
step: 510, loss: 0.001173494616523385
step: 520, loss: 4.963577157468535e-05
step: 530, loss: 9.532836702419445e-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=0.9426811937470393, f1=0.9289461134954697,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.7512675994075835e-05
step: 10, loss: 3.5080996894976124e-05
step: 20, loss: 2.3297296138480306e-05
step: 30, loss: 4.002577770734206e-05
step: 40, loss: 0.00031219847733154893
step: 50, loss: 5.338179107639007e-05
step: 60, loss: 0.000122987839858979
step: 70, loss: 5.7718949392437935e-05
step: 80, loss: 3.188712435076013e-05
step: 90, loss: 0.004347348120063543
step: 100, loss: 4.382743645692244e-05
step: 110, loss: 2.4354825654882006e-05
step: 120, loss: 4.899547275272198e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 3.069498052354902e-05
step: 140, loss: 0.024998169392347336
step: 150, loss: 3.937036308343522e-05
step: 160, loss: 2.9100816391292028e-05
step: 170, loss: 0.00018553962581790984
step: 180, loss: 3.965747600886971e-05
step: 190, loss: 5.4899956012377515e-05
step: 200, loss: 0.0009075568523257971
step: 210, loss: 7.589806045871228e-05
step: 220, loss: 0.0018675640458241105
step: 230, loss: 0.0005760424537584186

```

```

step: 240, loss: 5.074880027677864e-05
step: 250, loss: 0.01158139854669571
step: 260, loss: 0.005940528586506844
step: 270, loss: 8.713200077181682e-05
step: 280, loss: 4.1657480323920026e-05
step: 290, loss: 2.3591346689499915e-05
step: 300, loss: 4.2568473872961476e-05
step: 310, loss: 4.728779458673671e-05
step: 320, loss: 0.001180183608084917
step: 330, loss: 7.966346311150119e-05
step: 340, loss: 0.0034244321286678314
step: 350, loss: 0.00030332576716318727
step: 360, loss: 0.00023101277474779636
step: 370, loss: 0.0007305177277885377
step: 380, loss: 5.291961497277953e-05
step: 390, loss: 8.527279715053737e-05
step: 400, loss: 0.000789826619438827
step: 410, loss: 0.002624299144372344
step: 420, loss: 0.002082743216305971
step: 430, loss: 2.7614451028057374e-05
step: 440, loss: 0.00011336176248732954
step: 450, loss: 6.462175952037796e-05
step: 460, loss: 0.0007641255506314337
step: 470, loss: 0.0006042382447049022
step: 480, loss: 5.072398926131427e-05
step: 490, loss: 0.00010019500041380525
step: 500, loss: 0.018155846744775772
step: 510, loss: 2.4001550627872348e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 520, loss: 0.0001117128340411
step: 530, loss: 0.0005976061802357435
/usr/local/lib/python3.7/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=0.9438727782974743, f1=0.9338959212376934,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045857459190301597
step: 10, loss: 0.00014811534492764622
step: 20, loss: 9.53650232986547e-05
step: 30, loss: 6.986211519688368e-05
step: 40, loss: 2.7283171220915392e-05
step: 50, loss: 5.157927080290392e-05
step: 60, loss: 0.033714376389980316
step: 70, loss: 5.617092392640188e-05
step: 80, loss: 0.0034565983805805445

```

```
step: 90, loss: 3.462400854914449e-05
step: 100, loss: 5.889629028388299e-05
step: 110, loss: 5.5183023505378515e-05
step: 120, loss: 3.1382969609694555e-05
step: 130, loss: 1.7128564650192857e-05
step: 140, loss: 0.0003572425921447575
step: 150, loss: 2.0380070054670796e-05
step: 160, loss: 0.0006565955700352788
step: 170, loss: 2.1461026335600764e-05
step: 180, loss: 0.08646046370267868
step: 190, loss: 0.00045711526763625443
step: 200, loss: 0.011105363257229328
step: 210, loss: 7.529578579124063e-05
step: 220, loss: 0.0008810049621388316
step: 230, loss: 1.4442766769207083e-05
step: 240, loss: 7.809117960277945e-05
step: 250, loss: 2.3211108782561496e-05
step: 260, loss: 0.00024588051019236445
step: 270, loss: 4.765731137013063e-05
step: 280, loss: 6.926063360879198e-05
step: 290, loss: 4.3452899262774736e-05
step: 300, loss: 3.3916643587872386e-05
step: 310, loss: 4.166589860687964e-05
step: 320, loss: 5.179561412660405e-05
step: 330, loss: 0.0006813897634856403
step: 340, loss: 0.000214895058888942
step: 350, loss: 0.009772783145308495
step: 360, loss: 0.00020906038116663694
step: 370, loss: 0.00016544928075745702
step: 380, loss: 2.6381598217994906e-05
step: 390, loss: 0.0008095899829640985
step: 400, loss: 0.003339992370456457
step: 410, loss: 7.809314411133528e-05
step: 420, loss: 0.0011539481347426772
step: 430, loss: 4.2507817852310836e-05
step: 440, loss: 0.00020141474669799209
step: 450, loss: 5.1165530749130994e-05
step: 460, loss: 0.0019923895597457886
step: 470, loss: 4.9260896048508584e-05
step: 480, loss: 0.00017841071530710906
step: 490, loss: 4.839798930333927e-05
step: 500, loss: 0.0001993340119952336
step: 510, loss: 5.9225421864539385e-05
step: 520, loss: 3.277388896094635e-05
step: 530, loss: 5.3297600970836356e-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=0.9448818897637795, f1=0.9343880874825501,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.238218050682917e-05
step: 10, loss: 1.9248127500759438e-05
step: 20, loss: 0.00027829923783428967
step: 30, loss: 0.00016800445155240595
step: 40, loss: 4.8731315473560244e-05
step: 50, loss: 0.03876360505819321
step: 60, loss: 0.0006514628184959292
step: 70, loss: 9.005928586702794e-05
step: 80, loss: 1.702422378002666e-05
step: 90, loss: 0.0002716962480917573
step: 100, loss: 0.0001754845434334129
step: 110, loss: 0.020859306678175926
step: 120, loss: 2.7072963348473422e-05
step: 130, loss: 0.00033770507434383035
step: 140, loss: 1.9743622033274733e-05
step: 150, loss: 1.80710521817673e-05
step: 160, loss: 1.4874883163429331e-05
step: 170, loss: 5.900898759136908e-05
step: 180, loss: 1.9605802663136274e-05
step: 190, loss: 9.44972489378415e-05
step: 200, loss: 8.221752796089277e-05
step: 210, loss: 4.6133405703585595e-05
step: 220, loss: 3.610367639339529e-05
step: 230, loss: 1.8693041056394577e-05
step: 240, loss: 4.258898843545467e-05
step: 250, loss: 1.6264375517494045e-05
step: 260, loss: 2.21872851398075e-05
step: 270, loss: 3.926609133486636e-05
step: 280, loss: 0.000943221733905375
step: 290, loss: 0.00021182688942644745
step: 300, loss: 0.000177851426997222
step: 310, loss: 3.653501335065812e-05
step: 320, loss: 2.289028088853229e-05
step: 330, loss: 2.818779285007622e-05
step: 340, loss: 1.861124110291712e-05
step: 350, loss: 0.02379649691283703
step: 360, loss: 0.0029041613452136517
step: 370, loss: 0.0003043978940695524
step: 380, loss: 2.664628300408367e-05
step: 390, loss: 3.2232521334663033e-05
step: 400, loss: 0.00016678436077199876
step: 410, loss: 0.0004341963503975421
step: 420, loss: 0.004352975636720657
```



```

step: 430, loss: 0.02151818387210369
step: 440, loss: 2.502096504031215e-05
step: 450, loss: 3.52097958966624e-05
step: 460, loss: 1.5035072465252597e-05
step: 470, loss: 1.7631289665587246e-05
step: 480, loss: 6.897660932736471e-05
step: 490, loss: 0.0012686975533142686
step: 500, loss: 2.552439946157392e-05
step: 510, loss: 0.0030434392392635345
step: 520, loss: 1.856279413914308e-05
step: 530, loss: 3.996132727479562e-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=0.9400660689004247, f1=0.9291784702549575,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003095507563557476
step: 10, loss: 2.9747310691163875e-05
step: 20, loss: 1.1164574971189722e-05
step: 30, loss: 2.7070918804383837e-05
step: 40, loss: 0.00011581056605791673
step: 50, loss: 4.328288559918292e-05
step: 60, loss: 2.2171932869241573e-05
step: 70, loss: 2.894535100494977e-05
step: 80, loss: 4.4796681322623044e-05
step: 90, loss: 0.021398313343524933
step: 100, loss: 0.0030702606309205294
step: 110, loss: 2.3069529561325908e-05
step: 120, loss: 1.7437598216929473e-05
step: 130, loss: 6.523258343804628e-05
step: 140, loss: 3.858354830299504e-05
step: 150, loss: 4.9545451474841684e-05
step: 160, loss: 1.6197234799619764e-05
step: 170, loss: 0.0002226006326964125
step: 180, loss: 2.358411620662082e-05
step: 190, loss: 5.1850947784259915e-05
step: 200, loss: 2.5825793272815645e-05
step: 210, loss: 1.9162420358043164e-05
step: 220, loss: 1.596633592271246e-05
step: 230, loss: 1.3928686712461058e-05
step: 240, loss: 0.00035443840897642076
step: 250, loss: 1.7530759578221478e-05
step: 260, loss: 1.599979259481188e-05
step: 270, loss: 1.4181931874190923e-05
step: 280, loss: 2.009342824749183e-05

```

```

step: 290, loss: 2.1374480638769455e-05
step: 300, loss: 2.3874641556176357e-05
step: 310, loss: 0.00020016742928419262
step: 320, loss: 0.006255873944610357
step: 330, loss: 0.0016806027851998806
step: 340, loss: 2.7051337383454666e-05
step: 350, loss: 0.0010236725211143494
step: 360, loss: 1.6629448509775102e-05
step: 370, loss: 4.042845466756262e-05
step: 380, loss: 4.0360799175687134e-05
step: 390, loss: 0.09070850163698196
step: 400, loss: 0.00029714847914874554
step: 410, loss: 5.4334010201273486e-05
step: 420, loss: 1.737443744787015e-05
step: 430, loss: 0.00046761185512878
step: 440, loss: 0.001139962114393711
step: 450, loss: 9.286971180699766e-05
step: 460, loss: 0.0006453404785133898
step: 470, loss: 0.0001286948681809008
step: 480, loss: 1.8763787011266686e-05
step: 490, loss: 1.8488250134396367e-05
step: 500, loss: 1.2051190424244851e-05
step: 510, loss: 1.3682800272363238e-05
step: 520, loss: 0.00021514076797757298
step: 530, loss: 1.4450200978899375e-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=0.9443920829406222, f1=0.9331436699857754,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011975310189882293
step: 10, loss: 1.4315966836875305e-05
step: 20, loss: 1.4513559108308982e-05
step: 30, loss: 3.908440703526139e-05
step: 40, loss: 2.3580298147862777e-05
step: 50, loss: 9.137632150668651e-05
step: 60, loss: 1.3191046491556335e-05
step: 70, loss: 3.760116669582203e-05
step: 80, loss: 2.2887243176228367e-05
step: 90, loss: 2.2283822545432486e-05
step: 100, loss: 0.00010690849740058184
step: 110, loss: 1.7921929611475207e-05
step: 120, loss: 2.7326541385264136e-05
step: 130, loss: 0.012445134110748768
step: 140, loss: 0.00019349827198311687

```

```

step: 150, loss: 1.1481221008580178e-05
step: 160, loss: 2.2492415155284107e-05
step: 170, loss: 1.930412145156879e-05
step: 180, loss: 1.4982903849158902e-05
step: 190, loss: 1.205864919029409e-05
step: 200, loss: 2.299931838933844e-05
step: 210, loss: 1.5102134057087824e-05
step: 220, loss: 1.4472580915025901e-05
step: 230, loss: 2.32076217798749e-05
step: 240, loss: 2.9366103262873366e-05
step: 250, loss: 3.335319706820883e-05
step: 260, loss: 3.692528116516769e-05
step: 270, loss: 1.9661452824948356e-05
step: 280, loss: 1.311286905547604e-05
step: 290, loss: 9.639983181841671e-05
step: 300, loss: 2.3915612473501824e-05
step: 310, loss: 2.8124413802288473e-05
step: 320, loss: 7.113030005712062e-05
step: 330, loss: 3.348408790770918e-05
step: 340, loss: 1.5042525774333626e-05
step: 350, loss: 2.353556737944018e-05
step: 360, loss: 4.393558265292086e-05
step: 370, loss: 1.773946132743731e-05
step: 380, loss: 2.108839908032678e-05
step: 390, loss: 0.027293361723423004
step: 400, loss: 2.279441468999721e-05
step: 410, loss: 1.2930314369441476e-05
step: 420, loss: 1.4509724678646307e-05
step: 430, loss: 2.873563425964676e-05
step: 440, loss: 4.077706398675218e-05
step: 450, loss: 3.68363325833343e-05
step: 460, loss: 6.125698564574122e-05
step: 470, loss: 1.4185728105076123e-05
step: 480, loss: 2.7139914891449735e-05
step: 490, loss: 2.127771222149022e-05
step: 500, loss: 1.2155496733612381e-05
step: 510, loss: 2.035072429862339e-05
step: 520, loss: 2.3419312128680758e-05
step: 530, loss: 0.00018600348266772926
/usr/local/lib/python3.7/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=0.9453860640301318, f1=0.9323237103644108,
best_f1=0.9419475655430711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.935970431077294e-05

```

step: 10, loss: 4.869093754678033e-05
step: 20, loss: 2.485636832716409e-05
step: 30, loss: 1.6524978491361253e-05
step: 40, loss: 0.017031798139214516
step: 50, loss: 0.00019365940534044057
step: 60, loss: 1.2162908205937129e-05
step: 70, loss: 0.0002518514811526984
step: 80, loss: 1.830185647122562e-05
step: 90, loss: 1.7478767404099926e-05
step: 100, loss: 0.00012602057540789247
step: 110, loss: 2.4481458240188658e-05
step: 120, loss: 0.0038132192566990852
step: 130, loss: 0.0448867492377758
step: 140, loss: 1.4569397535524331e-05
step: 150, loss: 2.228416815341916e-05
step: 160, loss: 1.667786273173988e-05
step: 170, loss: 1.2717988283839077e-05
step: 180, loss: 1.012526081467513e-05
step: 190, loss: 1.7307460439042188e-05
step: 200, loss: 3.529457171680406e-05
step: 210, loss: 1.998539846681524e-05
step: 220, loss: 1.3824265806761105e-05
step: 230, loss: 1.6387260984629393e-05
step: 240, loss: 2.7601507099461742e-05
step: 250, loss: 3.0627452360931784e-05
step: 260, loss: 2.0957728338544257e-05
step: 270, loss: 1.6122805391205475e-05
step: 280, loss: 2.1698577256756835e-05
step: 290, loss: 1.1529657967912499e-05
step: 300, loss: 1.334382250206545e-05
step: 310, loss: 0.001163122826255858
step: 320, loss: 0.0001788902300177142
step: 330, loss: 0.00010116826160810888
step: 340, loss: 1.164884633908514e-05
step: 350, loss: 3.5781249607680365e-05
step: 360, loss: 0.0004124814295209944
step: 370, loss: 1.3004698303120676e-05
step: 380, loss: 3.225011096219532e-05
step: 390, loss: 0.0011024258565157652
step: 400, loss: 2.4923916498664767e-05
step: 410, loss: 6.088752707000822e-05
step: 420, loss: 2.352445517317392e-05
step: 430, loss: 6.253285391721874e-05
step: 440, loss: 2.4727491108933464e-05
step: 450, loss: 1.2237452210683841e-05
step: 460, loss: 1.462902218918316e-05
step: 470, loss: 0.0028592857997864485
step: 480, loss: 1.1585517313505989e-05

```

step: 490, loss: 1.4029261365067214e-05
step: 500, loss: 2.7750513254432008e-05
step: 510, loss: 1.7117155948653817e-05
step: 520, loss: 1.2498226169554982e-05
step: 530, loss: 1.7232901882380247e-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=0.9457802923149459, f1=0.9341544291804832,
best_f1=0.9419475655430711
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, 260.93it/s]
load_f1 = 0.9470290188853063
real_f1 = 0.9491990846681923
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 261.74it/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: 0.5528045296669006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4210526315789473, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5068184733390808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5, f1=0.43137254901960786, best_f1=0.43137254901960786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5750608444213867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.5925925925925927, f1=0.5161290322580646,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2028067708015442
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7333333333333334, f1=0.48484848484848486,
best_f1=0.48484848484848486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13508279621601105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8888888888888889, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07655778527259827
/usr/local/lib/python3.7/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=0.75, f1=0.7407407407407408, best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06130422651767731
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.7027027027027025,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03852448612451553
/usr/local/lib/python3.7/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=0.7142857142857143, best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002777969464659691
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.7058823529411764,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.055151186883449554
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.7222222222222223,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003350530518218875
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.742857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005445719696581364
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.742857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005392719060182571
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.742857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003352808067575097
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.742857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012262329459190369
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.742857142857143,
best_f1=0.6666666666666666
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 133082.87it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.9285714285714286
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 [00:16, 266.45it/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: 0.5979942679405212
step: 10, loss: 0.6248906850814819
step: 20, loss: 0.3217616081237793
step: 30, loss: 0.1096407100558281
step: 40, loss: 0.21241165697574615
step: 50, loss: 0.06309989094734192
step: 60, loss: 0.03343517705798149
step: 70, loss: 0.12684845924377441
step: 80, loss: 0.05317792668938637
step: 90, loss: 0.10868614166975021
step: 100, loss: 0.004982171580195427
step: 110, loss: 0.09974990785121918
step: 120, loss: 0.005864313337951899
step: 130, loss: 0.009247392416000366
step: 140, loss: 0.0018089966615661979
step: 150, loss: 0.012413130141794682
step: 160, loss: 0.02056683972477913
step: 170, loss: 0.31008878350257874
step: 180, loss: 0.042750682681798935
step: 190, loss: 0.009266888722777367
step: 200, loss: 0.07328518480062485
step: 210, loss: 0.002462411532178521
step: 220, loss: 0.083402618765831
step: 230, loss: 0.1009770855307579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.983050847457627, f1=0.9783845278725825,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004085567779839039
step: 10, loss: 0.0027263702359050512
step: 20, loss: 0.056486956775188446
step: 30, loss: 0.19093191623687744
step: 40, loss: 0.020802106708288193
step: 50, loss: 0.0030546218622475863
step: 60, loss: 0.0008123709121719003
step: 70, loss: 0.16897962987422943
step: 80, loss: 0.0011151550570502877

```

```

step: 90, loss: 0.0037604800891131163
step: 100, loss: 0.025004154071211815
step: 110, loss: 0.050808653235435486
step: 120, loss: 0.001507999375462532
step: 130, loss: 0.002681164303794503
step: 140, loss: 0.0005814743344672024
step: 150, loss: 0.0011131609790027142
step: 160, loss: 0.011309347115457058
step: 170, loss: 0.0012255813926458359
step: 180, loss: 0.0247269868850708
step: 190, loss: 0.1378936618566513
step: 200, loss: 0.0021859174594283104
step: 210, loss: 0.003151891054585576
step: 220, loss: 0.03170863911509514
step: 230, loss: 0.008525550365447998
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887892376681614, f1=0.9865470852017937,
best_f1=0.9865470852017937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019583485554903746
step: 10, loss: 0.028408553451299667
step: 20, loss: 0.002949690679088235
step: 30, loss: 0.09097941219806671
step: 40, loss: 0.1754218190908432
step: 50, loss: 0.012553631328046322
step: 60, loss: 0.004502158612012863
step: 70, loss: 0.0033050873316824436
step: 80, loss: 0.001384903327561915
step: 90, loss: 0.11040642112493515
step: 100, loss: 0.0027139890007674694
step: 110, loss: 0.0016682724235579371
step: 120, loss: 0.009298065677285194
step: 130, loss: 0.03507193922996521
step: 140, loss: 0.0096597233787179
step: 150, loss: 0.02572132647037506
step: 160, loss: 0.020959705114364624
step: 170, loss: 0.005561186000704765
step: 180, loss: 0.03815913572907448
step: 190, loss: 0.007855231873691082
step: 200, loss: 0.03783036768436432

```

```

step: 210, loss: 0.0009539004531688988
step: 220, loss: 0.09509730339050293
step: 230, loss: 0.006610924378037453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910313901345291, f1=0.9843400447427293,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012899631401523948
step: 10, loss: 0.0005726227536797523
step: 20, loss: 0.00040958053432404995
step: 30, loss: 0.0005075322114862502
step: 40, loss: 0.003726270282641053
step: 50, loss: 0.021221082657575607
step: 60, loss: 0.0007480258354917169
step: 70, loss: 0.00045493419747799635
step: 80, loss: 0.018649091944098473
step: 90, loss: 0.0009484911570325494
step: 100, loss: 0.00024608190869912505
step: 110, loss: 0.1810551881790161
step: 120, loss: 0.05484011769294739
step: 130, loss: 0.007387593854218721
step: 140, loss: 0.0007785160560160875
step: 150, loss: 0.10074513405561447
step: 160, loss: 0.0653814747929573
step: 170, loss: 0.0182168111205101
step: 180, loss: 0.0006235442124307156
step: 190, loss: 0.008613222278654575
step: 200, loss: 0.0024422896094620228
step: 210, loss: 0.1986188292503357
step: 220, loss: 0.001069836551323533
step: 230, loss: 0.005165127106010914
/usr/local/lib/python3.7/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=0.9910313901345291, f1=0.9865470852017937,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011523209977895021

```

```

step: 10, loss: 0.001134719350375235
step: 20, loss: 0.0011016272474080324
step: 30, loss: 0.0014766429085284472
step: 40, loss: 0.00042870582547038794
step: 50, loss: 0.0001579967065481469
step: 60, loss: 0.0018331209430471063
step: 70, loss: 0.0006806250894442201
step: 80, loss: 0.0005779421771876514
step: 90, loss: 0.0006948462105356157
step: 100, loss: 0.00033105112379416823
step: 110, loss: 0.00018079299479722977
step: 120, loss: 0.00010353512334404513
step: 130, loss: 0.0009082518517971039
step: 140, loss: 0.0006041749729774892
step: 150, loss: 0.0005012897308915854
step: 160, loss: 0.0005472422926686704
step: 170, loss: 0.012097216211259365
step: 180, loss: 0.006338601466268301
step: 190, loss: 0.010376577265560627
step: 200, loss: 0.029578499495983124
step: 210, loss: 0.00029708456713706255
step: 220, loss: 0.00039987373747862875
step: 230, loss: 0.00028832617681473494
/usr/local/lib/python3.7/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=0.9818594104308391, f1=0.9819004524886877,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23385974764823914
step: 10, loss: 0.00032966872095130384
step: 20, loss: 0.0013921469217166305
step: 30, loss: 0.000982710043899715
step: 40, loss: 0.0006420579156838357
step: 50, loss: 0.0003457627317402512
step: 60, loss: 0.00038169854087755084
step: 70, loss: 0.0003313649503979832
step: 80, loss: 0.0005560741992667317
step: 90, loss: 0.00041241536382585764
step: 100, loss: 0.08371686190366745
step: 110, loss: 0.0007406497024931014
step: 120, loss: 0.0011649514781311154
step: 130, loss: 0.0011530667543411255
step: 140, loss: 0.00039400512469001114
step: 150, loss: 0.0005597946001216769
step: 160, loss: 0.0007017817115411162

```



```

step: 170, loss: 0.0005492462660185993
step: 180, loss: 0.030450237914919853
step: 190, loss: 0.0005256748991087079
step: 200, loss: 0.013794628903269768
step: 210, loss: 0.0025708277244120836
step: 220, loss: 0.00017296044097747654
step: 230, loss: 0.08039543032646179
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9854423292273236,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003513350966386497
step: 10, loss: 9.814182703848928e-05
step: 20, loss: 0.0005476144724525511
step: 30, loss: 0.002957196906208992
step: 40, loss: 0.0003574960574042052
step: 50, loss: 0.0002634350094012916
step: 60, loss: 0.0016905387165024877
step: 70, loss: 0.02907121554017067
step: 80, loss: 0.003134090220555663
step: 90, loss: 0.00017495288921054453
step: 100, loss: 0.0001732591917971149
step: 110, loss: 0.0012918499996885657
step: 120, loss: 0.0020995549857616425
step: 130, loss: 0.0005964362062513828
step: 140, loss: 0.00011901280959136784
step: 150, loss: 0.006918569561094046
step: 160, loss: 0.061506569385528564
step: 170, loss: 0.0029896004125475883
step: 180, loss: 0.0003696754283737391
step: 190, loss: 0.00012248613347765058
step: 200, loss: 0.039576441049575806
step: 210, loss: 7.667656609555706e-05
step: 220, loss: 0.00028406534693203866
step: 230, loss: 0.00014046789146959782
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9833147942157954,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.000408368185163e-05

```

```

step: 10, loss: 0.0057199876755476
step: 20, loss: 0.00029233883833512664
step: 30, loss: 0.00011217566498089582
step: 40, loss: 0.002681640675291419
step: 50, loss: 0.0004735269758384675
step: 60, loss: 0.00018916754925157875
step: 70, loss: 0.0004524465766735375
step: 80, loss: 0.00011664251360343769
step: 90, loss: 5.151218647370115e-05
step: 100, loss: 0.0007817771984264255
step: 110, loss: 0.18854530155658722
step: 120, loss: 0.00016056901949923486
step: 130, loss: 0.0001250399654963985
step: 140, loss: 0.0001388432428939268
step: 150, loss: 0.000660075806081295
step: 160, loss: 0.0003957643057219684
step: 170, loss: 0.00015313878247980028
step: 180, loss: 0.00010541195661062375
step: 190, loss: 0.00024897162802517414
step: 200, loss: 0.00010245717567158863
step: 210, loss: 0.0001957191270776093
step: 220, loss: 0.005844404920935631
step: 230, loss: 0.00010678699618438259
/usr/local/lib/python3.7/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=0.9899441340782122, f1=0.9800884955752212,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.559638859471306e-05
step: 10, loss: 0.00048738333862274885
step: 20, loss: 0.0011552078649401665
step: 30, loss: 0.0002799740177579224
step: 40, loss: 0.029101060703396797
step: 50, loss: 6.713135371683165e-05
step: 60, loss: 0.0007580722449347377
step: 70, loss: 0.0001984167902264744
step: 80, loss: 9.099271119339392e-05
step: 90, loss: 0.00012991830590181053
step: 100, loss: 0.00018018792616203427
step: 110, loss: 5.281416088109836e-05
step: 120, loss: 7.496462058043107e-05
step: 130, loss: 0.00023150350898504257
step: 140, loss: 6.591114652110264e-05
step: 150, loss: 0.0007819432648830116
step: 160, loss: 0.00012672012962866575

```

```

step: 170, loss: 0.00030988280195742846
step: 180, loss: 0.0005191686213947833
step: 190, loss: 8.927031740313396e-05
step: 200, loss: 4.566185816656798e-05
step: 210, loss: 0.000306733010802418
step: 220, loss: 5.385946496971883e-05
step: 230, loss: 0.00015522807370871305
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9865771812080537,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.816117820562795e-05
step: 10, loss: 2.4493141609127633e-05
step: 20, loss: 0.00012060698645655066
step: 30, loss: 0.00021880125859752297
step: 40, loss: 0.0011242353357374668
step: 50, loss: 0.00022458593593910336
step: 60, loss: 0.0001502063823863864
step: 70, loss: 0.0008222166216000915
step: 80, loss: 7.815256685717031e-05
step: 90, loss: 0.00014822503726463765
step: 100, loss: 0.00010653664503479376
step: 110, loss: 0.001169246039353311
step: 120, loss: 0.00026256442652083933
step: 130, loss: 0.00019574858015403152
step: 140, loss: 0.03781213238835335
step: 150, loss: 0.011944998055696487
step: 160, loss: 6.522744661197066e-05
step: 170, loss: 0.00021584084606729448
step: 180, loss: 0.0001665664167376235
step: 190, loss: 0.02501332014799118
step: 200, loss: 7.20540847396478e-05
step: 210, loss: 6.467047933256254e-05
step: 220, loss: 8.332843572134152e-05
step: 230, loss: 0.0009031220106408
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9876819708846584,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.8582456656731665e-05

```

```

step: 10, loss: 8.901517139747739e-05
step: 20, loss: 0.000226857271627523
step: 30, loss: 0.017781279981136322
step: 40, loss: 0.00015632080612704158
step: 50, loss: 0.002713036723434925
step: 60, loss: 0.0004025883099529892
step: 70, loss: 0.0005812860326841474
step: 80, loss: 2.8505262889666483e-05
step: 90, loss: 3.4963723010150716e-05
step: 100, loss: 6.1020553403068334e-05
step: 110, loss: 0.013388220220804214
step: 120, loss: 3.981432382715866e-05
step: 130, loss: 0.0009580188780091703
step: 140, loss: 0.016324549913406372
step: 150, loss: 0.027146048843860626
step: 160, loss: 4.290985452826135e-05
step: 170, loss: 0.021590137854218483
step: 180, loss: 8.408620487898588e-05
step: 190, loss: 4.9585905799176544e-05
step: 200, loss: 0.00014331525017041713
step: 210, loss: 2.623667023726739e-05
step: 220, loss: 4.788532532984391e-05
step: 230, loss: 0.0001378842571284622
/usr/local/lib/python3.7/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=0.987736900780379, f1=0.9822222222222222,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.333668276667595e-05
step: 10, loss: 3.884193938574754e-05
step: 20, loss: 5.098704787087627e-05
step: 30, loss: 0.0009846891043707728
step: 40, loss: 0.00020218304416630417
step: 50, loss: 0.0007789921946823597
step: 60, loss: 0.0013759697321802378
step: 70, loss: 3.25281907862518e-05
step: 80, loss: 5.7491361076245084e-05
step: 90, loss: 9.803642751649022e-05
step: 100, loss: 0.0001372037804685533
step: 110, loss: 0.00021217360335867852
step: 120, loss: 0.000191102983080782
step: 130, loss: 6.581656634807587e-05
step: 140, loss: 0.005700668785721064
step: 150, loss: 7.422481576213613e-05
step: 160, loss: 4.4686061301035807e-05

```

```

step: 170, loss: 7.880543853389099e-05
step: 180, loss: 0.08673439919948578
step: 190, loss: 0.00023257148859556764
step: 200, loss: 2.533498809498269e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 210, loss: 5.5647069530095905e-05
step: 220, loss: 0.0022486792877316475
step: 230, loss: 0.039851006120443344
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9865168539325843,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.534273704048246e-05
step: 10, loss: 0.0007182482513599098
step: 20, loss: 0.0011805705726146698
step: 30, loss: 8.462843834422529e-05
step: 40, loss: 6.700515950797126e-05
step: 50, loss: 0.00010680417472030967
step: 60, loss: 0.0002529826888348907
step: 70, loss: 4.965389598510228e-05
step: 80, loss: 7.098215428413823e-05
step: 90, loss: 0.00013224974100012332
step: 100, loss: 3.040492083528079e-05
step: 110, loss: 0.016092607751488686
step: 120, loss: 0.03330212086439133
step: 130, loss: 0.0026986568700522184
step: 140, loss: 9.14454140001908e-05
step: 150, loss: 0.00010101778025273234
step: 160, loss: 0.0017145891906693578
step: 170, loss: 0.0002857428044080734
step: 180, loss: 8.582918235333636e-05
step: 190, loss: 6.096352080930956e-05
step: 200, loss: 0.00010662865679478273
step: 210, loss: 3.443922469159588e-05
step: 220, loss: 4.5116372348275036e-05
step: 230, loss: 6.179293268360198e-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=0.988814317673378, f1=0.9843749999999999,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 4.393798008095473e-05
step: 10, loss: 0.015037653967738152
step: 20, loss: 0.000569905445445329
step: 30, loss: 0.0007814987329766154
step: 40, loss: 0.014173311181366444
step: 50, loss: 6.55288458801806e-05
step: 60, loss: 4.969928340869956e-05
step: 70, loss: 7.06634164089337e-05
step: 80, loss: 4.841797999688424e-05
step: 90, loss: 6.490949454018846e-05
step: 100, loss: 3.625333192758262e-05
step: 110, loss: 0.000949431152548641
step: 120, loss: 2.7860440241056494e-05
step: 130, loss: 5.8289686421630904e-05
step: 140, loss: 0.0001131652170442976
step: 150, loss: 4.010821430711076e-05
step: 160, loss: 0.005438098683953285
step: 170, loss: 4.7881039790809155e-05
step: 180, loss: 4.994032860849984e-05
step: 190, loss: 0.00010224484140053391
step: 200, loss: 5.902195698581636e-05
step: 210, loss: 0.00010316432599211112
step: 220, loss: 7.980441296240315e-05
step: 230, loss: 7.820110477041453e-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=0.9898989898989898, f1=0.9876543209876544,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.062902684789151e-05
step: 10, loss: 2.5119143174379133e-05
step: 20, loss: 4.29354295192752e-05
step: 30, loss: 6.700192898279056e-05
step: 40, loss: 0.0001580133248353377
step: 50, loss: 0.0520341731607914
step: 60, loss: 6.311175820883363e-05
step: 70, loss: 4.787803118233569e-05
step: 80, loss: 4.941925362800248e-05
step: 90, loss: 0.00014889669546391815
step: 100, loss: 5.54811667825561e-05
step: 110, loss: 2.979418059112504e-05
step: 120, loss: 0.0002789083810057491
step: 130, loss: 0.030466122552752495
step: 140, loss: 5.041633630753495e-05
step: 150, loss: 0.001742066233418882

```

```

step: 160, loss: 3.3366337447660044e-05
step: 170, loss: 3.444607864366844e-05
step: 180, loss: 0.00027414632495492697
step: 190, loss: 0.00037210978916846216
step: 200, loss: 0.00013997558562550694
step: 210, loss: 3.780678162002005e-05
step: 220, loss: 5.804312968393788e-05
step: 230, loss: 3.8084719562903047e-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=0.9843749999999999,
best_f1=0.9843400447427293
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 207.19it/s]
load_f1 = 0.9910313901345291
real_f1 = 0.9898989898989898
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, 249.65it/s]
```

4.1.3 DBLP-GoogleScholar - Running the matcher

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

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

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

```
Downloading: 100% 433/433 [00:00<00:00, 384kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 1.22MB/s]
Downloading: 100% 440M/440M [00:14<00:00, 30.2MB/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: 0.625394344329834
step: 10, loss: 0.5059318542480469
step: 20, loss: 0.47454124689102173
step: 30, loss: 0.08586743474006653
step: 40, loss: 0.14751656353473663
step: 50, loss: 0.1506640911102295
step: 60, loss: 0.07282403856515884
step: 70, loss: 0.058041639626026154
step: 80, loss: 0.049511004239320755
step: 90, loss: 0.17664098739624023
step: 100, loss: 0.01293342188000679
step: 110, loss: 0.13056418299674988
step: 120, loss: 0.029255283996462822
step: 130, loss: 0.06446830183267593
step: 140, loss: 0.06360827386379242
step: 150, loss: 0.02389485202729702

```

```

step: 160, loss: 0.0243091844022274
step: 170, loss: 0.0860695093870163
step: 180, loss: 0.08452426642179489
step: 190, loss: 0.0050957039929926395
step: 200, loss: 0.11223235726356506
step: 210, loss: 0.08455348759889603
step: 220, loss: 0.19352109730243683
step: 230, loss: 0.13486087322235107
step: 240, loss: 0.06795345991849899
step: 250, loss: 0.02706541121006012
step: 260, loss: 0.08374249190092087
step: 270, loss: 0.026634296402335167
step: 280, loss: 0.0210100244730711
step: 290, loss: 0.08140301704406738
step: 300, loss: 0.023435648530721664
step: 310, loss: 0.19572514295578003
step: 320, loss: 0.15415988862514496
step: 330, loss: 0.016448525711894035
step: 340, loss: 0.029267603531479836
step: 350, loss: 0.07024497538805008
step: 360, loss: 0.029669396579265594
step: 370, loss: 0.14442287385463715
step: 380, loss: 0.016536187380552292
step: 390, loss: 0.11527888476848602
step: 400, loss: 0.16190338134765625
step: 410, loss: 0.04395836591720581
step: 420, loss: 0.013913903385400772
step: 430, loss: 0.13096372783184052
step: 440, loss: 0.014269569888710976
step: 450, loss: 0.004410558845847845
step: 460, loss: 0.004887456074357033
step: 470, loss: 0.179748073220253
step: 480, loss: 0.059037674218416214
step: 490, loss: 0.042844418436288834
step: 500, loss: 0.066884346306324
step: 510, loss: 0.055964261293411255
step: 520, loss: 0.04628710448741913
step: 530, loss: 0.008566811680793762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9382379654859219, f1=0.9324932493249325,
best_f1=0.9324932493249325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.14230915904045105
step: 10, loss: 0.024434391409158707
step: 20, loss: 0.020011894404888153
step: 30, loss: 0.05440710484981537
step: 40, loss: 0.10539747029542923
step: 50, loss: 0.18868088722229004
step: 60, loss: 0.0031480579636991024
step: 70, loss: 0.026125382632017136
step: 80, loss: 0.021940456703305244
step: 90, loss: 0.012057175859808922
step: 100, loss: 0.013952868990600109
step: 110, loss: 0.039061956107616425
step: 120, loss: 0.02706928923726082
step: 130, loss: 0.054153334349393845
step: 140, loss: 0.015475970692932606
step: 150, loss: 0.07205268740653992
step: 160, loss: 0.009191152639687061
step: 170, loss: 0.014921462163329124
step: 180, loss: 0.03768441081047058
step: 190, loss: 0.04296634718775749
step: 200, loss: 0.003957650158554316
step: 210, loss: 0.11694090813398361
step: 220, loss: 0.05055994167923927
step: 230, loss: 0.0061779688112437725
step: 240, loss: 0.06514714658260345
step: 250, loss: 0.02557610347867012
step: 260, loss: 0.002854022430256009
step: 270, loss: 0.16878926753997803
step: 280, loss: 0.18072500824928284
step: 290, loss: 0.011187352240085602
step: 300, loss: 0.15841439366340637
step: 310, loss: 0.04241186007857323
step: 320, loss: 0.1209561750292778
step: 330, loss: 0.014861198142170906
step: 340, loss: 0.018680181354284286
step: 350, loss: 0.0023994550574570894
step: 360, loss: 0.053425416350364685
step: 370, loss: 0.09929586946964264
step: 380, loss: 0.05667145922780037
step: 390, loss: 0.027031738311052322
step: 400, loss: 0.08210581541061401
step: 410, loss: 0.010704750195145607
step: 420, loss: 0.024453727528452873
step: 430, loss: 0.10640585422515869
step: 440, loss: 0.09015952795743942
step: 450, loss: 0.025838162750005722
```

```

step: 460, loss: 0.14048834145069122
step: 470, loss: 0.13624705374240875
step: 480, loss: 0.17716406285762787
step: 490, loss: 0.007809016387909651
step: 500, loss: 0.09677563607692719
step: 510, loss: 0.005709343124181032
step: 520, loss: 0.03985929861664772
step: 530, loss: 0.007478676270693541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9449749886311961, f1=0.9411231884057971,
best_f1=0.9411231884057971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10538844019174576
step: 10, loss: 0.029947243630886078
step: 20, loss: 0.03795304894447327
step: 30, loss: 0.08202677965164185
step: 40, loss: 0.012247631326317787
step: 50, loss: 0.07420358806848526
step: 60, loss: 0.07486951351165771
step: 70, loss: 0.008665281347930431
step: 80, loss: 0.002033689059317112
step: 90, loss: 0.004483134485781193
step: 100, loss: 0.08562014251947403
step: 110, loss: 0.003165668807923794
step: 120, loss: 0.01343762967735529
step: 130, loss: 0.0018423881847411394
step: 140, loss: 0.07169552147388458
step: 150, loss: 0.01639677584171295
step: 160, loss: 0.008216442540287971
step: 170, loss: 0.051384586840867996
step: 180, loss: 0.006696073338389397
step: 190, loss: 0.008593994192779064
step: 200, loss: 0.0034546703100204468
step: 210, loss: 0.08898045122623444
step: 220, loss: 0.05564466863870621
step: 230, loss: 0.08014336228370667
step: 240, loss: 0.0050781359896063805
step: 250, loss: 0.06711520999670029
step: 260, loss: 0.013140030205249786
step: 270, loss: 0.002226345706731081

```

```

step: 280, loss: 0.0801745355129242
step: 290, loss: 0.0051216669380664825
step: 300, loss: 0.03449784591794014
step: 310, loss: 0.002184342360123992
step: 320, loss: 0.007674878928810358
step: 330, loss: 0.001961497589945793
step: 340, loss: 0.0015318038640543818
step: 350, loss: 0.07822337746620178
step: 360, loss: 0.019300511106848717
step: 370, loss: 0.05276941508054733
step: 380, loss: 0.020405881106853485
step: 390, loss: 0.004267930053174496
step: 400, loss: 0.01956179365515709
step: 410, loss: 0.002908918308094144
step: 420, loss: 0.15172836184501648
step: 430, loss: 0.015548262745141983
step: 440, loss: 0.01742575690150261
step: 450, loss: 0.07628623396158218
step: 460, loss: 0.04281558468937874
step: 470, loss: 0.17846332490444183
step: 480, loss: 0.0089803421869874
step: 490, loss: 0.003839803859591484
step: 500, loss: 0.008584053255617619
step: 510, loss: 0.04306693747639656
step: 520, loss: 0.05308295786380768
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 530, loss: 0.08234353363513947
/usr/local/lib/python3.7/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=0.9430670339761248, f1=0.9363468634686346,
best_f1=0.9411231884057971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02217697724699974
step: 10, loss: 0.028618713840842247
step: 20, loss: 0.012526883743703365
step: 30, loss: 0.0009993569692596793
step: 40, loss: 0.002674683928489685
step: 50, loss: 0.003066742094233632
step: 60, loss: 0.01506948471069336
step: 70, loss: 0.0026762739289551973
step: 80, loss: 0.004084238316863775
step: 90, loss: 0.03628503158688545
step: 100, loss: 0.006406046450138092
step: 110, loss: 0.007346093654632568
step: 120, loss: 0.0008363290689885616

```

```

step: 130, loss: 0.0016478594625368714
step: 140, loss: 0.0022278414107859135
step: 150, loss: 0.012121755629777908
step: 160, loss: 0.1119835153222084
step: 170, loss: 0.00981269869953394
step: 180, loss: 0.002802367089316249
step: 190, loss: 0.05688156187534332
step: 200, loss: 0.004399741068482399
step: 210, loss: 0.08723698556423187
step: 220, loss: 0.031282663345336914
step: 230, loss: 0.037996623665094376
step: 240, loss: 0.001443292130716145
step: 250, loss: 0.028018875047564507
step: 260, loss: 0.0011806018883362412
step: 270, loss: 0.005613280460238457
step: 280, loss: 0.04996519163250923
step: 290, loss: 0.027881542220711708
step: 300, loss: 0.00013030538684688509
step: 310, loss: 0.02275540865957737
step: 320, loss: 0.009783574379980564
step: 330, loss: 0.06838683784008026
step: 340, loss: 0.10252276808023453
step: 350, loss: 0.06368657201528549
step: 360, loss: 0.07737938314676285
step: 370, loss: 0.006824176758527756
step: 380, loss: 0.007954473607242107
step: 390, loss: 0.04527423903346062
step: 400, loss: 0.1820870190858841
step: 410, loss: 0.015870967879891396
step: 420, loss: 0.00443209009245038
step: 430, loss: 0.08873660862445831
step: 440, loss: 0.021575357764959335
step: 450, loss: 0.008078099228441715
step: 460, loss: 0.0009019632125273347
step: 470, loss: 0.053294070065021515
step: 480, loss: 0.12586411833763123
step: 490, loss: 0.006795862223953009
step: 500, loss: 0.028296422213315964
step: 510, loss: 0.01835642382502556
step: 520, loss: 0.08881998807191849
step: 530, loss: 0.09285616874694824
/usr/local/lib/python3.7/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=0.939825447864033, f1=0.9365586490187129,
best_f1=0.9411231884057971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07096506655216217
step: 10, loss: 0.04264390096068382
step: 20, loss: 0.0012008351041004062
step: 30, loss: 0.006016825791448355
step: 40, loss: 0.007229572627693415
step: 50, loss: 0.014451617375016212
step: 60, loss: 0.107066310942173
step: 70, loss: 0.022870011627674103
step: 80, loss: 0.0009436891414225101
step: 90, loss: 0.009678700938820839
step: 100, loss: 0.030086137354373932
step: 110, loss: 0.0002162037417292595
step: 120, loss: 0.0034066690132021904
step: 130, loss: 0.00024276568728964776
step: 140, loss: 0.0004481595824472606
step: 150, loss: 0.004941526334732771
step: 160, loss: 0.0008872015168890357
step: 170, loss: 0.012347518466413021
step: 180, loss: 0.0009289614390581846
step: 190, loss: 0.0005131234647706151
step: 200, loss: 0.00039725415990687907
step: 210, loss: 0.024253826588392258
step: 220, loss: 0.0032949326559901237
step: 230, loss: 0.01158925797790289
step: 240, loss: 0.019776422530412674
step: 250, loss: 0.00186646799556911
step: 260, loss: 0.018068093806505203
step: 270, loss: 0.0017251268727704883
step: 280, loss: 0.00474376929923892
step: 290, loss: 0.13124524056911469
step: 300, loss: 0.013057999312877655
step: 310, loss: 0.008705384097993374
step: 320, loss: 0.12095819413661957
step: 330, loss: 0.009017729200422764
step: 340, loss: 0.018821053206920624
step: 350, loss: 0.0015835391823202372
step: 360, loss: 0.035427846014499664
step: 370, loss: 0.006716176401823759
step: 380, loss: 0.00764639675617218
step: 390, loss: 0.001061324030160904
step: 400, loss: 0.0016963283997029066
step: 410, loss: 0.0006759561947546899
step: 420, loss: 0.0032027342822402716
step: 430, loss: 0.0017566144233569503
step: 440, loss: 0.010661466978490353
step: 450, loss: 0.008991307578980923
step: 460, loss: 0.0009700109367258847
```

```

step: 470, loss: 0.026369456201791763
step: 480, loss: 0.0026639802381396294
step: 490, loss: 0.0008540306007489562
step: 500, loss: 0.05676432326436043
step: 510, loss: 0.18156179785728455
step: 520, loss: 0.04081244766712189
step: 530, loss: 0.048724912106990814
/usr/local/lib/python3.7/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=0.9431123648330982, f1=0.9435897435897437,
best_f1=0.9411231884057971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018141252221539617
step: 10, loss: 0.00043975640437565744
step: 20, loss: 0.10366439074277878
step: 30, loss: 0.0015762801049277186
step: 40, loss: 0.06794705241918564
step: 50, loss: 0.026118692010641098
step: 60, loss: 0.0005071493797004223
step: 70, loss: 0.000577304104808718
step: 80, loss: 0.005099987145513296
step: 90, loss: 0.0001401044282829389
step: 100, loss: 0.0007648806786164641
step: 110, loss: 0.007599365431815386
step: 120, loss: 0.001098967739380896
step: 130, loss: 0.001488619134761393
step: 140, loss: 0.013519524596631527
step: 150, loss: 0.0010766308987513185
step: 160, loss: 0.018022671341896057
step: 170, loss: 0.036337029188871384
step: 180, loss: 0.00117104547098279
step: 190, loss: 0.0030556339770555496
step: 200, loss: 0.03159784525632858
step: 210, loss: 0.003071230137720704
step: 220, loss: 0.0009781972039490938
step: 230, loss: 0.009500709362328053
step: 240, loss: 0.07082200795412064
step: 250, loss: 0.025909030809998512
step: 260, loss: 0.006220785900950432
step: 270, loss: 0.0673031434416771
step: 280, loss: 0.006905817426741123
step: 290, loss: 0.0015528808580711484
step: 300, loss: 0.002484675729647279
step: 310, loss: 0.0009652996086515486
step: 320, loss: 0.00024884240701794624

```



```
step: 330, loss: 0.0002121751313097775
step: 340, loss: 0.06691113859415054
step: 350, loss: 0.019488098099827766
step: 360, loss: 0.023665431886911392
step: 370, loss: 0.003890320658683777
step: 380, loss: 0.0014546531019732356
step: 390, loss: 0.03606884554028511
step: 400, loss: 0.0020755836740136147
step: 410, loss: 0.00044342968612909317
step: 420, loss: 0.024011384695768356
step: 430, loss: 0.0021621224004775286
step: 440, loss: 0.0007495247409678996
step: 450, loss: 0.00037424336187541485
step: 460, loss: 0.00045035668881610036
step: 470, loss: 0.021458743140101433
step: 480, loss: 0.0006990641122683883
step: 490, loss: 0.011081207543611526
step: 500, loss: 0.0003247983695473522
step: 510, loss: 0.026547886431217194
step: 520, loss: 0.004149147309362888
step: 530, loss: 0.04887351766228676
/usr/local/lib/python3.7/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=0.9395218002812938, f1=0.9432029795158287,
best_f1=0.9411231884057971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001995197031646967
step: 10, loss: 0.001921714749187231
step: 20, loss: 0.00681092357262969
step: 30, loss: 0.0014991058269515634
step: 40, loss: 0.04138299077749252
step: 50, loss: 0.005915014538913965
step: 60, loss: 0.11137116700410843
step: 70, loss: 0.002112798159942031
step: 80, loss: 0.004441659897565842
step: 90, loss: 0.0005618691211566329
step: 100, loss: 0.004395924974232912
step: 110, loss: 0.001156953047029674
step: 120, loss: 0.002811664482578635
step: 130, loss: 0.006071142386645079
step: 140, loss: 0.0003934115229640156
step: 150, loss: 0.00908007100224495
step: 160, loss: 0.0001580162497702986
step: 170, loss: 0.00015028638881631196
step: 180, loss: 0.0183649230748415
```

```

step: 190, loss: 0.006486371625214815
step: 200, loss: 0.004769566003233194
step: 210, loss: 0.001994863385334611
step: 220, loss: 0.0009943487821146846
step: 230, loss: 0.08023914694786072
step: 240, loss: 0.00045178941218182445
step: 250, loss: 0.0035291037056595087
step: 260, loss: 0.022951534017920494
step: 270, loss: 0.006422777660191059
step: 280, loss: 0.007768433541059494
step: 290, loss: 0.000405343365855515
step: 300, loss: 0.0025235603097826242
step: 310, loss: 0.0006223914679139853
step: 320, loss: 0.000535168219357729
step: 330, loss: 0.04918641597032547
step: 340, loss: 0.024852901697158813
step: 350, loss: 0.050241194665431976
step: 360, loss: 0.0015964668709784746
step: 370, loss: 0.0005084769218228757
step: 380, loss: 0.004238035064190626
step: 390, loss: 0.00041278928983956575
step: 400, loss: 0.0038215487729758024
step: 410, loss: 0.0019206006545573473
step: 420, loss: 0.0005038303206674755
step: 430, loss: 0.00023592830984853208
step: 440, loss: 0.08359653502702713
step: 450, loss: 0.002662929240614176
step: 460, loss: 0.0016624617855995893
step: 470, loss: 0.0293581485748291
step: 480, loss: 0.09362654387950897
step: 490, loss: 0.0008172435918822885
step: 500, loss: 0.007546009495854378
step: 510, loss: 0.002600599080324173
step: 520, loss: 0.025678930804133415
step: 530, loss: 0.009831894189119339
/usr/local/lib/python3.7/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=0.9424131627056673, f1=0.9378427787934186,
best_f1=0.9411231884057971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003952756058424711
step: 10, loss: 0.019057540223002434
step: 20, loss: 0.02133430354297161
step: 30, loss: 0.0013290238566696644
step: 40, loss: 0.00011162841110490263

```

step: 50, loss: 0.0010750218061730266
step: 60, loss: 0.0013884907821193337
step: 70, loss: 0.0004351880925241858
step: 80, loss: 0.00034294891520403326
step: 90, loss: 0.003293949645012617
step: 100, loss: 0.0015516221756115556
step: 110, loss: 0.0007108912686817348
step: 120, loss: 0.01820867881178856
step: 130, loss: 0.0016993982717394829
step: 140, loss: 0.013382001779973507
step: 150, loss: 0.007211368530988693
step: 160, loss: 0.01947285234928131
step: 170, loss: 0.02690603770315647
step: 180, loss: 0.0011942654382437468
step: 190, loss: 0.0028659889940172434
step: 200, loss: 0.002507997676730156
step: 210, loss: 0.03665165975689888
step: 220, loss: 0.004975358955562115
step: 230, loss: 0.00047205400187522173
step: 240, loss: 0.00028319505508989096
step: 250, loss: 0.000797587854322046
step: 260, loss: 0.13321904838085175
step: 270, loss: 0.005660699680447578
step: 280, loss: 0.01605820097029209
step: 290, loss: 0.0004912505391985178
step: 300, loss: 0.019758600741624832
step: 310, loss: 0.00014928462042007595
step: 320, loss: 0.012576314620673656
step: 330, loss: 0.0005762993241660297
step: 340, loss: 0.00028310398920439184
step: 350, loss: 0.0012989743845537305
step: 360, loss: 0.0018380866385996342
step: 370, loss: 0.00045492115896195173
step: 380, loss: 0.0004947020206600428
step: 390, loss: 0.11962448060512543
step: 400, loss: 0.01322785671800375
step: 410, loss: 0.0001865123922470957
step: 420, loss: 0.002630977425724268
step: 430, loss: 0.0242423415184021
step: 440, loss: 0.0015358919044956565
step: 450, loss: 0.00024815762299112976
step: 460, loss: 0.001585829653777182
step: 470, loss: 0.0016662583220750093
step: 480, loss: 0.0002069606416625902
step: 490, loss: 0.0007276133983395994
step: 500, loss: 0.0001766929344739765
step: 510, loss: 0.05774984136223793
step: 520, loss: 0.0023877352941781282

```

step: 530, loss: 0.0003239240904804319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9470260223048327, f1=0.94362292051756,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002093295333907008
step: 10, loss: 0.0035113783087581396
step: 20, loss: 0.0012995621655136347
step: 30, loss: 0.002139780903235078
step: 40, loss: 0.00027564394986256957
step: 50, loss: 0.0002526501484680921
step: 60, loss: 0.0002694558061193675
step: 70, loss: 0.0015391606139019132
step: 80, loss: 0.0027466765604913235
step: 90, loss: 6.973450217628852e-05
step: 100, loss: 0.0003088365192525089
step: 110, loss: 0.00044575511128641665
step: 120, loss: 0.005657605826854706
step: 130, loss: 0.018439602106809616
step: 140, loss: 0.0009648416889831424
step: 150, loss: 0.00024158200540114194
step: 160, loss: 0.0012793955393135548
step: 170, loss: 0.015345558524131775
step: 180, loss: 0.0008697540615685284
step: 190, loss: 0.0008026389987207949
step: 200, loss: 0.00038206164026632905
step: 210, loss: 0.0017590246861800551
step: 220, loss: 0.002257788786664605
step: 230, loss: 0.00011763790098484606
step: 240, loss: 0.0005487306625582278
step: 250, loss: 0.0004808113444596529
step: 260, loss: 0.03787228837609291
step: 270, loss: 6.690133886877447e-05
step: 280, loss: 0.0015584152424708009
step: 290, loss: 0.04194993898272514
step: 300, loss: 0.00013486917305272073
step: 310, loss: 0.02045629546046257
step: 320, loss: 0.00021498753631021827
step: 330, loss: 0.000764252501539886
step: 340, loss: 9.42736878641881e-05

```

```
step: 350, loss: 0.005299811251461506
step: 360, loss: 0.0025893566198647022
step: 370, loss: 0.0018856028327718377
step: 380, loss: 0.00017994501104112715
step: 390, loss: 0.0005693556158803403
step: 400, loss: 0.00020352343562990427
step: 410, loss: 0.0019460071343928576
step: 420, loss: 3.536642179824412e-05
step: 430, loss: 0.0001552623143652454
step: 440, loss: 0.004242326598614454
step: 450, loss: 9.396391396876425e-05
step: 460, loss: 0.0006730480818077922
step: 470, loss: 3.3697888284223154e-05
step: 480, loss: 6.138004391686991e-05
step: 490, loss: 0.05073315277695656
step: 500, loss: 0.0017559295520186424
step: 510, loss: 0.00029954692581668496
step: 520, loss: 0.008268417790532112
step: 530, loss: 9.860943828243762e-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=0.9375293565054016, f1=0.9340813464235624,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2531737685203552
step: 10, loss: 0.00019678099488373846
step: 20, loss: 0.0006149844848550856
step: 30, loss: 0.007606710307300091
step: 40, loss: 0.002409488195553422
step: 50, loss: 0.00027896984829567373
step: 60, loss: 0.0003342948039062321
step: 70, loss: 0.003936955239623785
step: 80, loss: 0.000800839567091316
step: 90, loss: 0.1467345505952835
step: 100, loss: 0.0001573542394908145
step: 110, loss: 8.953525684773922e-05
step: 120, loss: 0.03782128915190697
step: 130, loss: 0.001171357580460608
step: 140, loss: 0.010269757360219955
step: 150, loss: 0.043135713785886765
step: 160, loss: 0.00038114713970571756
step: 170, loss: 0.0003725681744981557
step: 180, loss: 0.000915011391043663
step: 190, loss: 0.05043329671025276
step: 200, loss: 0.0012522959150373936
```

```

step: 210, loss: 0.0007005864754319191
step: 220, loss: 0.003874517511576414
step: 230, loss: 0.00594046525657177
step: 240, loss: 0.0016826011706143618
step: 250, loss: 0.00011884285777341574
step: 260, loss: 0.0016439097234979272
step: 270, loss: 0.0012817549286410213
step: 280, loss: 0.00015363082638941705
step: 290, loss: 0.0006933754193596542
step: 300, loss: 6.717291398672387e-05
step: 310, loss: 0.002961613703519106
step: 320, loss: 0.0216076597571373
step: 330, loss: 0.0005803977255709469
step: 340, loss: 0.012972088530659676
step: 350, loss: 0.0016744663007557392
step: 360, loss: 0.00038733409019187093
step: 370, loss: 0.001988789765164256
step: 380, loss: 0.09564666450023651
step: 390, loss: 0.00017731093976181
step: 400, loss: 0.0013074605958536267
step: 410, loss: 0.0020116614177823067
step: 420, loss: 0.002103522652760148
step: 430, loss: 0.00017363532970193774
step: 440, loss: 0.0019405840430408716
step: 450, loss: 0.00011382593220332637
step: 460, loss: 8.435647760052234e-05
step: 470, loss: 0.0002678585588000715
step: 480, loss: 0.0002945055312011391
step: 490, loss: 0.00024406574084423482
step: 500, loss: 0.012829157523810863
step: 510, loss: 0.000429064326453954
step: 520, loss: 0.00021011156786698848
step: 530, loss: 0.0009744334383867681
/usr/local/lib/python3.7/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=0.9445983379501385, f1=0.9425287356321841,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006013854406774044
step: 10, loss: 0.00255779130384326
step: 20, loss: 6.384232256095856e-05
step: 30, loss: 8.502630953444168e-05
step: 40, loss: 0.00013259411207400262
step: 50, loss: 0.010206964798271656
step: 60, loss: 0.03245801851153374

```

step: 70, loss: 8.948499453254044e-05
step: 80, loss: 0.0003584928344935179
step: 90, loss: 2.7272077204543166e-05
step: 100, loss: 0.0004341877356637269
step: 110, loss: 0.00023687677457928658
step: 120, loss: 0.00023428615531884134
step: 130, loss: 0.0006623392691835761
step: 140, loss: 0.001425790716893971
step: 150, loss: 5.403759496402927e-05
step: 160, loss: 3.780646875384264e-05
step: 170, loss: 0.00019853540288750082
step: 180, loss: 0.00011692992120515555
step: 190, loss: 0.002528655808418989
step: 200, loss: 0.005201494321227074
step: 210, loss: 0.00019824672199320048
step: 220, loss: 0.00135696679353714
step: 230, loss: 3.329130413476378e-05
step: 240, loss: 0.0002832534082699567
step: 250, loss: 0.0001916109031299129
step: 260, loss: 2.6742634872789495e-05
step: 270, loss: 8.412035822402686e-05
step: 280, loss: 3.4096163290087134e-05
step: 290, loss: 0.005663667339831591
step: 300, loss: 6.122361810412258e-05
step: 310, loss: 0.00011592222290346399
step: 320, loss: 8.063274435698986e-05
step: 330, loss: 0.0006562831695191562
step: 340, loss: 3.63712570106145e-05
step: 350, loss: 3.785045919357799e-05
step: 360, loss: 2.3871161829447374e-05
step: 370, loss: 0.001351657323539257
step: 380, loss: 0.00030448322650045156
step: 390, loss: 4.008114046882838e-05
step: 400, loss: 0.0006050942465662956
step: 410, loss: 0.0003181269858032465
step: 420, loss: 3.0508845156873576e-05
step: 430, loss: 8.512412750860676e-05
step: 440, loss: 0.00025841203751042485
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 450, loss: 0.01584695093333721
step: 460, loss: 0.06241970509290695
step: 470, loss: 6.382724677678198e-05
step: 480, loss: 0.0023551771882921457
step: 490, loss: 0.0008610765798948705
step: 500, loss: 8.836216147756204e-05
step: 510, loss: 0.0015724695986136794
step: 520, loss: 3.448698771535419e-05
step: 530, loss: 0.0002792786981444806

```
/usr/local/lib/python3.7/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=0.9409541804440246, f1=0.9382022471910113,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014018705405760556
step: 10, loss: 0.0001513142924522981
step: 20, loss: 0.002024248940870166
step: 30, loss: 0.0025731860660016537
step: 40, loss: 0.0011606367770582438
step: 50, loss: 0.05037693679332733
step: 60, loss: 0.002932630479335785
step: 70, loss: 0.0002976919349748641
step: 80, loss: 0.00013857471640221775
step: 90, loss: 0.00018689224089030176
step: 100, loss: 6.540687900269404e-05
step: 110, loss: 4.287251431378536e-05
step: 120, loss: 5.3830237447982654e-05
step: 130, loss: 0.0012221187353134155
step: 140, loss: 8.094054646790028e-05
step: 150, loss: 7.896440365584567e-05
step: 160, loss: 5.847972715855576e-05
step: 170, loss: 6.580797344213352e-05
step: 180, loss: 0.00016676404629833996
step: 190, loss: 0.00010701428982429206
step: 200, loss: 0.00019184098346158862
step: 210, loss: 0.0009254051838070154
step: 220, loss: 0.0009163996437564492
step: 230, loss: 5.713663267670199e-05
step: 240, loss: 0.0017420247895643115
step: 250, loss: 4.283722228137776e-05
step: 260, loss: 0.00014399809879250824
step: 270, loss: 5.5922242609085515e-05
step: 280, loss: 0.0011619420256465673
step: 290, loss: 0.0010619533713907003
step: 300, loss: 0.0010957479244098067
step: 310, loss: 0.00013356289127841592
step: 320, loss: 7.041732169454917e-05
step: 330, loss: 0.0024245474487543106
step: 340, loss: 0.00021315310732461512
step: 350, loss: 0.0027392141055315733
step: 360, loss: 0.0033904523588716984
step: 370, loss: 6.715233757859096e-05
step: 380, loss: 0.00043549443944357336
step: 390, loss: 0.00030631618574261665
```



```

step: 400, loss: 0.0030501573346555233
step: 410, loss: 0.001018040580675006
step: 420, loss: 0.01702699065208435
step: 430, loss: 0.02743157185614109
step: 440, loss: 0.012502382509410381
step: 450, loss: 0.0005948444595560431
step: 460, loss: 0.0001114367478294298
step: 470, loss: 4.434525908436626e-05
step: 480, loss: 0.00047196808736771345
step: 490, loss: 0.0005985988536849618
step: 500, loss: 6.555143045261502e-05
step: 510, loss: 0.016197780147194862
step: 520, loss: 2.7547741410671733e-05
step: 530, loss: 7.061436190269887e-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=0.9457579972183587, f1=0.9461573861021628,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001866710837930441
step: 10, loss: 5.118511398904957e-05
step: 20, loss: 5.1086048188153654e-05
step: 30, loss: 7.070192805258557e-05
step: 40, loss: 8.256710134446621e-05
step: 50, loss: 0.0049345968291163445
step: 60, loss: 0.0025943657383322716
step: 70, loss: 3.0753730243304744e-05
step: 80, loss: 0.00024989794474095106
step: 90, loss: 0.03494339436292648
step: 100, loss: 0.0008721204358153045
step: 110, loss: 0.02563413418829441
step: 120, loss: 0.001977639738470316
step: 130, loss: 0.0003245163825340569
step: 140, loss: 0.00023114957730285823
step: 150, loss: 0.0009269036236219108
step: 160, loss: 5.7308865507366136e-05
step: 170, loss: 0.0007801190949976444
step: 180, loss: 0.0009246512199752033
step: 190, loss: 0.0003184794040862471
step: 200, loss: 6.406768079614267e-05
step: 210, loss: 0.00011008106230292469
step: 220, loss: 0.0003802375285886228
step: 230, loss: 0.00013169576413929462
step: 240, loss: 0.00010774117254186422
step: 250, loss: 3.081800969084725e-05

```

```

step: 260, loss: 8.758119656704366e-05
step: 270, loss: 3.41951526934281e-05
step: 280, loss: 7.215191726572812e-05
step: 290, loss: 0.004461455624550581
step: 300, loss: 0.0020510961767286062
step: 310, loss: 0.0008142044534906745
step: 320, loss: 0.00457457872107625
step: 330, loss: 0.00392509438097477
step: 340, loss: 0.0006789880571886897
step: 350, loss: 0.004964679013937712
step: 360, loss: 0.0003016131231561303
step: 370, loss: 0.00023011054145172238
step: 380, loss: 5.198044891585596e-05
step: 390, loss: 0.0028371328953653574
step: 400, loss: 0.00020456660422496498
step: 410, loss: 0.0031616396736353636
step: 420, loss: 0.00022449219250120223
step: 430, loss: 0.007031257729977369
step: 440, loss: 0.1424260437488556
step: 450, loss: 0.00042779507930390537
step: 460, loss: 0.003941099159419537
step: 470, loss: 0.0001617837988305837
step: 480, loss: 0.0007516514742746949
step: 490, loss: 0.0005511263734661043
step: 500, loss: 0.0002515113155823201
step: 510, loss: 0.0017770014237612486
step: 520, loss: 0.002242538845166564
step: 530, loss: 0.00012196854368085042
/usr/local/lib/python3.7/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=0.9433255269320844, f1=0.9452181987000929,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010682581923902035
step: 10, loss: 4.3128828110639006e-05
step: 20, loss: 0.00013649002357851714
step: 30, loss: 3.8599529943894595e-05
step: 40, loss: 0.0005695111467503011
step: 50, loss: 0.0001820025936467573
step: 60, loss: 3.717436266015284e-05
step: 70, loss: 6.607710383832455e-05
step: 80, loss: 4.601979526341893e-05
step: 90, loss: 0.00011000357335433364
step: 100, loss: 0.00013961564400233328
step: 110, loss: 6.117430893937126e-05

```

```
step: 120, loss: 0.00012024625175399706
step: 130, loss: 0.012896613217890263
step: 140, loss: 0.0009522518375888467
step: 150, loss: 0.00013497043983079493
step: 160, loss: 0.0004091514565516263
step: 170, loss: 0.0010631862096488476
step: 180, loss: 9.798017708817497e-05
step: 190, loss: 0.001760778483003378
step: 200, loss: 0.04400813579559326
step: 210, loss: 0.0001339769223704934
step: 220, loss: 5.84163426537998e-05
step: 230, loss: 0.0005541307618841529
step: 240, loss: 0.0011592069640755653
step: 250, loss: 0.00025261365226469934
step: 260, loss: 0.0017542729619890451
step: 270, loss: 0.002245398238301277
step: 280, loss: 0.00012227907427586615
step: 290, loss: 0.0018171286210417747
step: 300, loss: 0.00022804398031439632
step: 310, loss: 0.0015780861722305417
step: 320, loss: 0.0010779383592307568
step: 330, loss: 0.00016591587336733937
step: 340, loss: 0.0007980309310369194
step: 350, loss: 0.0005824977997690439
step: 360, loss: 0.07323791831731796
step: 370, loss: 0.00011814671597676352
step: 380, loss: 0.00019677638192661107
step: 390, loss: 0.03234582394361496
step: 400, loss: 0.001099518034607172
step: 410, loss: 6.595869490411133e-05
step: 420, loss: 0.0008204157929867506
step: 430, loss: 0.0001029149498208426
step: 440, loss: 0.027309957891702652
step: 450, loss: 0.0023141526617109776
step: 460, loss: 0.0008039003587327898
step: 470, loss: 5.90464478591457e-05
step: 480, loss: 7.75927328504622e-05
step: 490, loss: 7.78109606471844e-05
step: 500, loss: 8.600152796134353e-05
step: 510, loss: 0.0016409209929406643
step: 520, loss: 0.0012002203147858381
step: 530, loss: 0.0043130372650921345
/usr/local/lib/python3.7/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=0.9439035697728326, f1=0.9478541762805722,
best_f1=0.94362292051756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

step: 0, loss: 0.0002523211878724396
step: 10, loss: 5.982038419460878e-05
step: 20, loss: 0.0019800576847046614
step: 30, loss: 0.00012591220729518682
step: 40, loss: 0.08730977773666382
step: 50, loss: 0.03356008976697922
step: 60, loss: 0.0001163773049484007
step: 70, loss: 0.0015801506815478206
step: 80, loss: 0.00021148164523765445
step: 90, loss: 6.534068961627781e-05
step: 100, loss: 0.00011100108531536534
step: 110, loss: 0.0017434563487768173
step: 120, loss: 0.00302168563939631
step: 130, loss: 0.11188112944364548
step: 140, loss: 0.00010972219752147794
step: 150, loss: 0.0014601628063246608
step: 160, loss: 0.00017006759298965335
step: 170, loss: 0.0001053761225193739
step: 180, loss: 5.069074177299626e-05
step: 190, loss: 4.1935305489460006e-05
step: 200, loss: 0.00014134650700725615
step: 210, loss: 0.0021295968908816576
step: 220, loss: 0.00010207825835095719
step: 230, loss: 0.00018406858725938946
step: 240, loss: 0.00018559132877271622
step: 250, loss: 8.06367170298472e-05
step: 260, loss: 0.0002961491118185222
step: 270, loss: 2.088725523208268e-05
step: 280, loss: 2.68546227744082e-05
step: 290, loss: 2.396051240793895e-05
step: 300, loss: 3.451343945926055e-05
step: 310, loss: 0.00013407898950390518
step: 320, loss: 0.0010163808474317193
step: 330, loss: 3.624930832302198e-05
step: 340, loss: 8.706939115654677e-05
step: 350, loss: 4.046952017233707e-05
step: 360, loss: 0.0015210799174383283
step: 370, loss: 0.0003178375191055238
step: 380, loss: 0.00010725442552939057
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 390, loss: 0.001942059607245028
step: 400, loss: 0.00012155037256889045
step: 410, loss: 0.00013522310473490506
step: 420, loss: 0.0002833815524354577
step: 430, loss: 0.0001579890085849911
step: 440, loss: 0.00019370135851204395

```

step: 450, loss: 2.022057378781028e-05
step: 460, loss: 0.001102912938222289
step: 470, loss: 0.0012594029540196061
step: 480, loss: 4.853271093452349e-05
step: 490, loss: 1.9363718820386566e-05
step: 500, loss: 9.365262667415664e-05
step: 510, loss: 4.406747029861435e-05
step: 520, loss: 3.988917887909338e-05
step: 530, loss: 0.00031241492251865566
/usr/local/lib/python3.7/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=0.9445471349353051, f1=0.9462068965517241,
best_f1=0.94362292051756
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 [00:25, 224.65it/s]
load_f1 = 0.9440591770688858
real_f1 = 0.9447004608294931
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 232.56it/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: 0.5684472322463989
step: 10, loss: 0.35652703046798706
step: 20, loss: 0.39123088121414185
step: 30, loss: 0.3232665956020355
step: 40, loss: 0.1653166562318802
step: 50, loss: 0.4122862219810486
step: 60, loss: 0.19825780391693115
step: 70, loss: 0.14340722560882568
step: 80, loss: 0.21056416630744934
step: 90, loss: 0.2897893786430359
step: 100, loss: 0.359599769115448
step: 110, loss: 0.20040340721607208
step: 120, loss: 0.19733931124210358
step: 130, loss: 0.1470964401960373
step: 140, loss: 0.21915581822395325
step: 150, loss: 0.24182969331741333
step: 160, loss: 0.3305444121360779
step: 170, loss: 0.24173569679260254

```

```

step: 180, loss: 0.09951762855052948
step: 190, loss: 0.1770179122686386
step: 200, loss: 0.24501347541809082
step: 210, loss: 0.2260449379682541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6168582375478926, f1=0.6627450980392158,
best_f1=0.6627450980392158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1089334711432457
step: 10, loss: 0.14277136325836182
step: 20, loss: 0.20940326154232025
step: 30, loss: 0.22021429240703583
step: 40, loss: 0.1674620807170868
step: 50, loss: 0.12102262675762177
step: 60, loss: 0.3674778640270233
step: 70, loss: 0.10486620664596558
step: 80, loss: 0.1799599528312683
step: 90, loss: 0.11546376347541809
step: 100, loss: 0.023988988250494003
step: 110, loss: 0.14257319271564484
step: 120, loss: 0.23505699634552002
step: 130, loss: 0.009786821901798248
step: 140, loss: 0.19973674416542053
step: 150, loss: 0.2080608308315277
step: 160, loss: 0.25875380635261536
step: 170, loss: 0.1682187020778656
step: 180, loss: 0.2551259994506836
step: 190, loss: 0.2517484724521637
step: 200, loss: 0.05882783606648445
step: 210, loss: 0.18148061633110046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.676, f1=0.7330677290836652, best_f1=0.7330677290836652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05228773504495621
step: 10, loss: 0.14292851090431213
step: 20, loss: 0.14313329756259918
step: 30, loss: 0.16892074048519135
step: 40, loss: 0.16102324426174164
step: 50, loss: 0.0871867686510086
step: 60, loss: 0.11365899443626404
step: 70, loss: 0.11763112246990204
step: 80, loss: 0.21348437666893005
step: 90, loss: 0.03821319714188576
step: 100, loss: 0.17790770530700684
step: 110, loss: 0.15793153643608093
step: 120, loss: 0.1348523646593094
step: 130, loss: 0.2435423731803894
step: 140, loss: 0.1300952434539795
step: 150, loss: 0.2149265557527542
step: 160, loss: 0.013030152767896652
step: 170, loss: 0.06612297892570496
step: 180, loss: 0.18168379366397858
step: 190, loss: 0.1783459633588791
step: 200, loss: 0.04762210324406624
step: 210, loss: 0.11608652025461197
/usr/local/lib/python3.7/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=0.6653465346534654, f1=0.70703125, best_f1=0.7330677290836652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10620339214801788
step: 10, loss: 0.04578199237585068
step: 20, loss: 0.015140200965106487
step: 30, loss: 0.06606833636760712
step: 40, loss: 0.005141433794051409
step: 50, loss: 0.19338448345661163
step: 60, loss: 0.188102126121521
step: 70, loss: 0.1459452360868454
step: 80, loss: 0.01575526036322117
step: 90, loss: 0.008565470576286316
step: 100, loss: 0.20771650969982147
step: 110, loss: 0.17783257365226746
step: 120, loss: 0.0894261971116066
step: 130, loss: 0.25296750664711
step: 140, loss: 0.09588607400655746
step: 150, loss: 0.12982851266860962
step: 160, loss: 0.0760893002152443
step: 170, loss: 0.14321269094944

```

```

step: 180, loss: 0.7878934741020203
step: 190, loss: 0.08897221088409424
step: 200, loss: 0.13299410045146942
step: 210, loss: 0.11251142621040344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6804511278195489, f1=0.7132075471698114,
best_f1=0.7132075471698114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08902812004089355
step: 10, loss: 0.08362994343042374
step: 20, loss: 0.22847695648670197
step: 30, loss: 0.06515920907258987
step: 40, loss: 0.12376227229833603
step: 50, loss: 0.09268327802419662
step: 60, loss: 0.10585226118564606
step: 70, loss: 0.13772915303707123
step: 80, loss: 0.0266402754932642
step: 90, loss: 0.26031601428985596
step: 100, loss: 0.007667133118957281
step: 110, loss: 0.12972144782543182
step: 120, loss: 0.11086087673902512
step: 130, loss: 0.04295526444911957
step: 140, loss: 0.1058254987001419
step: 150, loss: 0.03201476112008095
step: 160, loss: 0.11807219684123993
step: 170, loss: 0.050280649214982986
step: 180, loss: 0.06677268445491791
step: 190, loss: 0.011967462487518787
step: 200, loss: 0.09821800142526627
step: 210, loss: 0.03559475392103195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6956521739130436, f1=0.704, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06233227625489235
step: 10, loss: 0.0966743752360344
step: 20, loss: 0.08325672894716263
step: 30, loss: 0.015352596528828144
step: 40, loss: 0.01621941663324833
step: 50, loss: 0.02771790139377117
step: 60, loss: 0.13366833329200745
step: 70, loss: 0.07006079703569412
step: 80, loss: 0.17572994530200958
step: 90, loss: 0.07139886915683746
step: 100, loss: 0.01786971278488636
step: 110, loss: 0.059164080768823624
step: 120, loss: 0.1626262664794922
step: 130, loss: 0.05694320797920227
step: 140, loss: 0.12103137373924255
step: 150, loss: 0.07831839472055435
step: 160, loss: 0.0386635921895504
step: 170, loss: 0.11700336635112762
step: 180, loss: 0.019928984344005585
step: 190, loss: 0.07113053649663925
step: 200, loss: 0.006515795830637217
step: 210, loss: 0.04621646925806999
/usr/local/lib/python3.7/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=0.6872586872586872, f1=0.7114624505928853, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006159595213830471
step: 10, loss: 0.008056159131228924
step: 20, loss: 0.13623574376106262
step: 30, loss: 0.020951971411705017
step: 40, loss: 0.005097553133964539
step: 50, loss: 0.11808368563652039
step: 60, loss: 0.06083119660615921
step: 70, loss: 0.011323044076561928
step: 80, loss: 0.08602257817983627
step: 90, loss: 0.06564773619174957
step: 100, loss: 0.010957692749798298
step: 110, loss: 0.16994605958461761
step: 120, loss: 0.18234960734844208
step: 130, loss: 0.05310634523630142
step: 140, loss: 0.04012380167841911
step: 150, loss: 0.0583246685564518
step: 160, loss: 0.06565017253160477
step: 170, loss: 0.1980629712343216

```

```

step: 180, loss: 0.12338660657405853
step: 190, loss: 0.1248709037899971
step: 200, loss: 0.012106246314942837
step: 210, loss: 0.2725011110305786
/usr/local/lib/python3.7/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=0.6827309236947792, f1=0.6918489065606361, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.038910701870918274
step: 10, loss: 0.09651362895965576
step: 20, loss: 0.010163119994103909
step: 30, loss: 0.03720158338546753
step: 40, loss: 0.024328405037522316
step: 50, loss: 0.02553064562380314
step: 60, loss: 0.25189775228500366
step: 70, loss: 0.053488004952669144
step: 80, loss: 0.024372262880206108
step: 90, loss: 0.011220455169677734
step: 100, loss: 0.15915465354919434
step: 110, loss: 0.06135428324341774
step: 120, loss: 0.008350628428161144
step: 130, loss: 0.02964152954518795
step: 140, loss: 0.06167340651154518
step: 150, loss: 0.06151806563138962
step: 160, loss: 0.04180223494768143
step: 170, loss: 0.015515672042965889
step: 180, loss: 0.07547002285718918
step: 190, loss: 0.04262072965502739
step: 200, loss: 0.002999645657837391
step: 210, loss: 0.30589956045150757
/usr/local/lib/python3.7/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=0.6720977596741344, f1=0.6736842105263158, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029601216316223145
step: 10, loss: 0.05543017014861107
step: 20, loss: 0.009084544144570827
step: 30, loss: 0.06735227257013321
step: 40, loss: 0.037727102637290955
step: 50, loss: 0.031134385615587234
step: 60, loss: 0.10245770961046219
step: 70, loss: 0.011300170794129372

```

```

step: 80, loss: 0.003797123208642006
step: 90, loss: 0.002946353517472744
step: 100, loss: 0.031063620001077652
step: 110, loss: 0.04021096229553223
step: 120, loss: 0.02487097680568695
step: 130, loss: 0.0037290588952600956
step: 140, loss: 0.021767746657133102
step: 150, loss: 0.05818622186779976
step: 160, loss: 0.005757140461355448
step: 170, loss: 0.05484453961253166
step: 180, loss: 0.09357203543186188
step: 190, loss: 0.03181998431682587
step: 200, loss: 0.09219805896282196
step: 210, loss: 0.014352633617818356
/usr/local/lib/python3.7/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=0.6912065439672802, f1=0.6804979253112032, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02567223645746708
step: 10, loss: 0.046400267630815506
step: 20, loss: 0.0007874885923229158
step: 30, loss: 0.09203845262527466
step: 40, loss: 0.012088035233318806
step: 50, loss: 0.01503329910337925
step: 60, loss: 0.038858652114868164
step: 70, loss: 0.09883318096399307
step: 80, loss: 0.1330488622188568
step: 90, loss: 0.020692605525255203
step: 100, loss: 0.02576550841331482
step: 110, loss: 0.08229804784059525
step: 120, loss: 0.11029131710529327
step: 130, loss: 0.1530018299818039
step: 140, loss: 0.014248956926167011
step: 150, loss: 0.07538825273513794
step: 160, loss: 0.02519097737967968
step: 170, loss: 0.07267200201749802
step: 180, loss: 0.059513624757528305
step: 190, loss: 0.188935324549675
step: 200, loss: 0.01857227459549904
step: 210, loss: 0.07317060977220535
/usr/local/lib/python3.7/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=0.6716981132075471, f1=0.6755218216318785, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024297164753079414
step: 10, loss: 0.0411674901843071
step: 20, loss: 0.06857812404632568
step: 30, loss: 0.0009069163934327662
step: 40, loss: 0.027856498956680298
step: 50, loss: 0.01377352699637413
step: 60, loss: 0.026740437373518944
step: 70, loss: 0.06981153041124344
step: 80, loss: 0.04312439635396004
step: 90, loss: 0.03426305204629898
step: 100, loss: 0.015108569525182247
step: 110, loss: 0.009954629465937614
step: 120, loss: 0.06502681225538254
step: 130, loss: 0.013242412358522415
step: 140, loss: 0.015265988186001778
step: 150, loss: 0.04426579177379608
step: 160, loss: 0.0029591135680675507
step: 170, loss: 0.006916161626577377
step: 180, loss: 0.006175389513373375
step: 190, loss: 0.01530087273567915
step: 200, loss: 0.0028845882043242455
step: 210, loss: 0.003120200475677848
/usr/local/lib/python3.7/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=0.6728971962616823, f1=0.6914498141263941, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05824630334973335
step: 10, loss: 0.0010968622518703341
step: 20, loss: 0.05144056677818298
step: 30, loss: 0.0021515695843845606
step: 40, loss: 0.04001830145716667
step: 50, loss: 0.015880310907959938
step: 60, loss: 0.23009414970874786
step: 70, loss: 0.00326304673217237
step: 80, loss: 0.005017880816012621
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 90, loss: 0.15320807695388794
step: 100, loss: 0.026126710698008537
step: 110, loss: 0.015691664069890976
step: 120, loss: 0.03906361013650894
step: 130, loss: 0.0422804020345211
step: 140, loss: 0.002238831715658307
step: 150, loss: 0.03416779637336731

```

```

step: 160, loss: 0.0010315118124708533
step: 170, loss: 0.008744570426642895
step: 180, loss: 0.045154210180044174
step: 190, loss: 0.007184560410678387
step: 200, loss: 0.11415573209524155
step: 210, loss: 0.05090700834989548
/usr/local/lib/python3.7/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=0.6736401673640167, f1=0.676470588235294, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.050415486097335815
step: 10, loss: 0.0026323653291910887
step: 20, loss: 0.027465246617794037
step: 30, loss: 0.037272047251462936
step: 40, loss: 0.034919433295726776
step: 50, loss: 0.005204509478062391
step: 60, loss: 0.006218201946467161
step: 70, loss: 0.04821982979774475
step: 80, loss: 0.011273471638560295
step: 90, loss: 0.0009222417720593512
step: 100, loss: 0.0016059044282883406
step: 110, loss: 0.0037065462674945593
step: 120, loss: 0.0025251873303204775
step: 130, loss: 0.0011942057171836495
step: 140, loss: 0.14439767599105835
step: 150, loss: 0.0052510336972773075
step: 160, loss: 0.0029282893519848585
step: 170, loss: 0.002703688573092222
step: 180, loss: 0.045570120215415955
step: 190, loss: 0.007062319200485945
step: 200, loss: 0.005011531990021467
step: 210, loss: 0.02397027611732483
/usr/local/lib/python3.7/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=0.6756756756756757, f1=0.6933333333333335, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006105617620050907
step: 10, loss: 0.037143073976039886
step: 20, loss: 0.002783970208838582
step: 30, loss: 0.05381392315030098
step: 40, loss: 0.004728574305772781
step: 50, loss: 0.015748171135783195

```

```

step: 60, loss: 0.009970361366868019
step: 70, loss: 0.0022658153902739286
step: 80, loss: 0.002453066408634186
step: 90, loss: 0.18041929602622986
step: 100, loss: 0.013999405317008495
step: 110, loss: 0.0013511758297681808
step: 120, loss: 0.05731583386659622
step: 130, loss: 0.14360763132572174
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.021892953664064407
step: 150, loss: 0.018021181225776672
step: 160, loss: 0.02420269511640072
step: 170, loss: 0.002939244732260704
step: 180, loss: 0.055798497051000595
step: 190, loss: 0.007012493908405304
step: 200, loss: 0.026535963639616966
step: 210, loss: 0.041554711759090424
/usr/local/lib/python3.7/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=0.6796116504854368, f1=0.6986564299424184, best_f1=0.704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033874879591166973
step: 10, loss: 0.03169641271233559
step: 20, loss: 0.012638573534786701
step: 30, loss: 0.20404364168643951
step: 40, loss: 0.0019236556254327297
step: 50, loss: 0.0007766244816593826
step: 60, loss: 0.04307997226715088
step: 70, loss: 0.003436111146584153
step: 80, loss: 0.004248610697686672
step: 90, loss: 0.018850058317184448
step: 100, loss: 0.03751189261674881
step: 110, loss: 0.0036711590364575386
step: 120, loss: 0.010307222604751587
step: 130, loss: 0.02030419372022152
step: 140, loss: 0.0021060837898403406
step: 150, loss: 0.0015703109093010426
step: 160, loss: 0.004521846305578947
step: 170, loss: 0.0010529292048886418
step: 180, loss: 0.000979896984063089
step: 190, loss: 0.0014028584118932486
step: 200, loss: 0.00641379551962018
step: 210, loss: 0.06423734873533249
/usr/local/lib/python3.7/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=0.6816326530612244, f1=0.7004048582995952, best_f1=0.704
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, 415.55it/s]
load_f1 = 0.6872727272727273
real_f1 = 0.6838235294117647
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 [00:19, 229.98it/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: 0.5362911224365234
step: 10, loss: 0.38411280512809753
step: 20, loss: 0.2890129089355469
step: 30, loss: 0.4384654760360718
step: 40, loss: 0.42827168107032776
step: 50, loss: 0.2992008626461029
step: 60, loss: 0.2612677812576294
step: 70, loss: 0.28971561789512634
step: 80, loss: 0.24161937832832336
step: 90, loss: 0.2490578293800354
step: 100, loss: 0.3250889778137207
step: 110, loss: 0.40760406851768494
step: 120, loss: 0.0834604948759079
step: 130, loss: 0.12319968640804291
step: 140, loss: 0.03892624005675316
step: 150, loss: 0.1572478860616684
step: 160, loss: 0.12293311208486557
step: 170, loss: 0.29287225008010864
step: 180, loss: 0.027259251102805138
step: 190, loss: 0.1848987340927124
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7002652519893899, f1=0.7376623376623376,
best_f1=0.7376623376623376

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3683188855648041
step: 10, loss: 0.09444015473127365
step: 20, loss: 0.07978207617998123
step: 30, loss: 0.14730879664421082
step: 40, loss: 0.142686128616333
step: 50, loss: 0.04642019420862198
step: 60, loss: 0.18922339379787445
step: 70, loss: 0.2277245819568634
step: 80, loss: 0.14650097489356995
step: 90, loss: 0.2359740287065506
step: 100, loss: 0.24008236825466156
step: 110, loss: 0.2617310881614685
step: 120, loss: 0.21572642028331757
step: 130, loss: 0.043291497975587845
step: 140, loss: 0.0890645757317543
step: 150, loss: 0.05541089177131653
step: 160, loss: 0.033101797103881836
step: 170, loss: 0.041192956268787384
step: 180, loss: 0.11698175221681595
step: 190, loss: 0.016534317284822464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7641025641025642, f1=0.772378516624041,
best_f1=0.772378516624041
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07117511332035065
step: 10, loss: 0.14433175325393677
step: 20, loss: 0.12931296229362488
step: 30, loss: 0.021763814613223076
step: 40, loss: 0.03842092305421829
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.2660810351371765
step: 60, loss: 0.01801191084086895
step: 70, loss: 0.05309059098362923
step: 80, loss: 0.05535389110445976
step: 90, loss: 0.08106838166713715
step: 100, loss: 0.040121711790561676
step: 110, loss: 0.011215094476938248

```

```

step: 120, loss: 0.08884090185165405
step: 130, loss: 0.035201482474803925
step: 140, loss: 0.009065228514373302
step: 150, loss: 0.17388516664505005
step: 160, loss: 0.025600461289286613
step: 170, loss: 0.145729199051857
step: 180, loss: 0.02246711589396
step: 190, loss: 0.013036465272307396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8060453400503779, f1=0.7918781725888324,
best_f1=0.7918781725888324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040411606431007385
step: 10, loss: 0.09964954107999802
step: 20, loss: 0.05328894779086113
step: 30, loss: 0.005688859615474939
step: 40, loss: 0.02160003036260605
step: 50, loss: 0.014113491401076317
step: 60, loss: 0.12408068776130676
step: 70, loss: 0.12271968275308609
step: 80, loss: 0.2162468135356903
step: 90, loss: 0.0033461730927228928
step: 100, loss: 0.16583389043807983
step: 110, loss: 0.014875280670821667
step: 120, loss: 0.05830150842666626
step: 130, loss: 0.14715354144573212
step: 140, loss: 0.03837292641401291
step: 150, loss: 0.06613223999738693
step: 160, loss: 0.019612284377217293
step: 170, loss: 0.0646686851978302
step: 180, loss: 0.01990828849375248
step: 190, loss: 0.1926666647195816
/usr/local/lib/python3.7/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=0.8, f1=0.7771739130434783, best_f1=0.7918781725888324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053609699010849

```

```

step: 10, loss: 0.01954673044383526
step: 20, loss: 0.03350977599620819
step: 30, loss: 0.0020768423564732075
step: 40, loss: 0.021805549040436745
step: 50, loss: 0.19345691800117493
step: 60, loss: 0.03735063225030899
step: 70, loss: 0.008604603819549084
step: 80, loss: 0.01172056794166565
step: 90, loss: 0.02604611963033676
step: 100, loss: 0.007684539537876844
step: 110, loss: 0.13212862610816956
step: 120, loss: 0.004637476056814194
step: 130, loss: 0.16986259818077087
step: 140, loss: 0.01336275227367878
step: 150, loss: 0.09770473092794418
step: 160, loss: 0.026499267667531967
step: 170, loss: 0.07689131796360016
step: 180, loss: 0.0016038308385759592
step: 190, loss: 0.04297502338886261
/usr/local/lib/python3.7/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=0.8042328042328043, f1=0.8105263157894735,
best_f1=0.7918781725888324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00768816564232111
step: 10, loss: 0.005136644467711449
step: 20, loss: 0.006038586609065533
step: 30, loss: 0.004884315654635429
step: 40, loss: 0.109476238489151
step: 50, loss: 0.01288316398859024
step: 60, loss: 0.013877011835575104
step: 70, loss: 0.05307852104306221
step: 80, loss: 0.008147645741701126
step: 90, loss: 0.007814791984856129
step: 100, loss: 0.0013907020911574364
step: 110, loss: 0.06877648830413818
step: 120, loss: 0.0030811969190835953
step: 130, loss: 0.010726336389780045
step: 140, loss: 0.05087221413850784
step: 150, loss: 0.01723909005522728
step: 160, loss: 0.10352523624897003
step: 170, loss: 0.03305215761065483
step: 180, loss: 0.012553293257951736
step: 190, loss: 0.017287634313106537
/usr/local/lib/python3.7/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=0.7902439024390243, f1=0.7860696517412935,
best_f1=0.7918781725888324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005305446218699217
step: 10, loss: 0.01853019930422306
step: 20, loss: 0.031649842858314514
step: 30, loss: 0.0020004508551210165
step: 40, loss: 0.003016664879396558
step: 50, loss: 0.11231253296136856
step: 60, loss: 0.03597144037485123
step: 70, loss: 0.002031890908256173
step: 80, loss: 0.023168858140707016
step: 90, loss: 0.04779401421546936
step: 100, loss: 0.0002909470349550247
step: 110, loss: 0.0736628919839859
step: 120, loss: 0.0059614940546453
step: 130, loss: 0.0011160981375724077
step: 140, loss: 0.04458086937665939
step: 150, loss: 0.0027613951824605465
step: 160, loss: 0.07755549997091293
step: 170, loss: 0.1205226257443428
step: 180, loss: 0.0021294066682457924
step: 190, loss: 0.004856658633798361
/usr/local/lib/python3.7/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=0.8059701492537312, f1=0.8070175438596492,
best_f1=0.7918781725888324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006524091586470604
step: 10, loss: 0.04028291627764702
step: 20, loss: 0.010136591270565987
step: 30, loss: 0.02360207960009575
step: 40, loss: 0.002929063281044364
step: 50, loss: 0.0019503558287397027
step: 60, loss: 0.03341792896389961
step: 70, loss: 0.024219676852226257
step: 80, loss: 0.00025949414703063667
step: 90, loss: 0.009515506215393543
step: 100, loss: 0.008390510454773903
step: 110, loss: 0.0006360994302667677
step: 120, loss: 0.00784518662840128

```

```

step: 130, loss: 0.0010027341777458787
step: 140, loss: 0.001232560840435326
step: 150, loss: 0.006340373307466507
step: 160, loss: 0.005868491716682911
step: 170, loss: 0.0026585611049085855
step: 180, loss: 0.09091944992542267
step: 190, loss: 0.06718804687261581
/usr/local/lib/python3.7/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=0.7945205479452055, f1=0.8219178082191781,
best_f1=0.7918781725888324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0051566604524850845
step: 10, loss: 0.0673084482550621
step: 20, loss: 0.010237090289592743
step: 30, loss: 0.0003251034941058606
step: 40, loss: 0.0013704252196475863
step: 50, loss: 0.006305903661996126
step: 60, loss: 0.0008166849729605019
step: 70, loss: 0.030234308913350105
step: 80, loss: 0.00023838943161536008
step: 90, loss: 0.009834679774940014
step: 100, loss: 0.06253285706043243
step: 110, loss: 0.0016977888299152255
step: 120, loss: 0.0035518514923751354
step: 130, loss: 0.031441580504179
step: 140, loss: 0.04793081432580948
step: 150, loss: 0.0058525120839476585
step: 160, loss: 0.0010581318056210876
step: 170, loss: 0.0036410049069672823
step: 180, loss: 0.10827326029539108
step: 190, loss: 0.0011983021395280957
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8153846153846155, f1=0.8268733850129197,
best_f1=0.8268733850129197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002002198249101639

```



```

step: 10, loss: 0.0049604978412389755
step: 20, loss: 0.006832274608314037
step: 30, loss: 0.014818492345511913
step: 40, loss: 0.007271083537489176
step: 50, loss: 0.0019914000295102596
step: 60, loss: 0.04005748778581619
step: 70, loss: 0.006125019397586584
step: 80, loss: 0.00550270639359951
step: 90, loss: 0.013596116565167904
step: 100, loss: 0.0009718266082927585
step: 110, loss: 0.006431568879634142
step: 120, loss: 0.09579383581876755
step: 130, loss: 0.06828475743532181
step: 140, loss: 0.0020606517791748047
step: 150, loss: 0.0031883851625025272
step: 160, loss: 0.007820113562047482
step: 170, loss: 0.0016173400217667222
step: 180, loss: 0.0009047645726241171
step: 190, loss: 0.0038513904437422752
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8177083333333333, f1=0.8315789473684211,
best_f1=0.8315789473684211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000594812969211489
step: 10, loss: 0.0005161039298400283
step: 20, loss: 0.009881340898573399
step: 30, loss: 0.017046982422471046
step: 40, loss: 0.002717111026868224
step: 50, loss: 0.0023594589438289404
step: 60, loss: 0.0010603078408166766
step: 70, loss: 0.0006847640033811331
step: 80, loss: 0.00034075987059623003
step: 90, loss: 0.0005872030160389841
step: 100, loss: 0.00034463725751265883
step: 110, loss: 0.0006290845340117812
step: 120, loss: 0.003581134369596839
step: 130, loss: 0.00022919217008166015
step: 140, loss: 0.003872399451211095
step: 150, loss: 0.00018062458548229188
step: 160, loss: 0.000947505293879658

```

```

step: 170, loss: 0.002046749694272876
step: 180, loss: 0.0415041409432888
step: 190, loss: 0.00039534905226901174
/usr/local/lib/python3.7/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=0.8054054054054054, f1=0.8021390374331551,
best_f1=0.8315789473684211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010374977719038725
step: 10, loss: 0.008112352341413498
step: 20, loss: 0.0006000997964292765
step: 30, loss: 0.023175761103630066
step: 40, loss: 0.004491650965064764
step: 50, loss: 0.02878548577427864
step: 60, loss: 0.0007354563567787409
step: 70, loss: 0.019389377906918526
step: 80, loss: 0.0007997335051186383
step: 90, loss: 0.0006253221072256565
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.0012448553461581469
step: 110, loss: 0.0013254255754873157
step: 120, loss: 0.026614531874656677
step: 130, loss: 0.0017316797748208046
step: 140, loss: 0.0014823414385318756
step: 150, loss: 0.0005052703199908137
step: 160, loss: 0.0019329481292515993
step: 170, loss: 0.002688656561076641
step: 180, loss: 0.0007781184976920485
step: 190, loss: 0.002226517302915454
/usr/local/lib/python3.7/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=0.7927461139896373, f1=0.8, best_f1=0.8315789473684211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006163741461932659
step: 10, loss: 0.004027611576020718
step: 20, loss: 0.0008814005414023995
step: 30, loss: 0.025092700496315956
step: 40, loss: 0.0052671064622700214
step: 50, loss: 0.008264080621302128
step: 60, loss: 0.002195071429014206
step: 70, loss: 0.003438886720687151
step: 80, loss: 0.0008087596506811678

```

```

step: 90, loss: 0.001192365656606853
step: 100, loss: 0.0007884703809395432
step: 110, loss: 0.016159934923052788
step: 120, loss: 0.000567577953916043
step: 130, loss: 0.0005756985046900809
step: 140, loss: 0.0006432444788515568
step: 150, loss: 0.0005037093651480973
step: 160, loss: 0.0007172988262027502
step: 170, loss: 0.0008629545336589217
step: 180, loss: 0.000757077126763761
step: 190, loss: 0.005087309516966343
/usr/local/lib/python3.7/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=0.8071065989847714, f1=0.8102564102564102,
best_f1=0.8315789473684211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1085047721862793
step: 10, loss: 0.0006037647253833711
step: 20, loss: 0.0004997647483833134
step: 30, loss: 0.0004961509839631617
step: 40, loss: 0.007965032942593098
step: 50, loss: 0.0003494491393212229
step: 60, loss: 0.0025214552879333496
step: 70, loss: 0.000562207424081862
step: 80, loss: 0.00037421914748847485
step: 90, loss: 0.019431032240390778
step: 100, loss: 0.005583966616541147
step: 110, loss: 0.002853327663615346
step: 120, loss: 0.002907652873545885
step: 130, loss: 0.0013350896770134568
step: 140, loss: 0.0008276599692180753
step: 150, loss: 0.0005169041687622666
step: 160, loss: 0.007645050995051861
step: 170, loss: 0.02033117413520813
step: 180, loss: 0.0008777562179602683
step: 190, loss: 0.0011574182426556945
/usr/local/lib/python3.7/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=0.8063660477453581, f1=0.8177083333333333,
best_f1=0.8315789473684211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05137919634580612

```

```

step: 10, loss: 0.011441214941442013
step: 20, loss: 0.0009544030181132257
step: 30, loss: 0.0012015120591968298
step: 40, loss: 0.0002038402744801715
step: 50, loss: 0.0038237334229052067
step: 60, loss: 0.0012942702742293477
step: 70, loss: 0.00972114410251379
step: 80, loss: 0.08427020907402039
step: 90, loss: 0.0010476093739271164
step: 100, loss: 0.0004046909452881664
step: 110, loss: 0.00035675158142112195
step: 120, loss: 0.00071162567473948
step: 130, loss: 0.00025561710936017334
step: 140, loss: 0.04006711021065712
step: 150, loss: 0.0007539867656305432
step: 160, loss: 0.0066537875682115555
step: 170, loss: 0.0008401293307542801
step: 180, loss: 0.0006971709080971777
step: 190, loss: 0.00027328537544235587
/usr/local/lib/python3.7/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=0.8050632911392406, f1=0.8153846153846155,
best_f1=0.8315789473684211
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, 210.07it/s]
load_f1 = 0.7493261455525607
real_f1 = 0.7180851063829788
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:19, 223.34it/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: 0.6425092220306396
step: 10, loss: 0.380376398563385
step: 20, loss: 0.31414735317230225
step: 30, loss: 0.3710547983646393
step: 40, loss: 0.28108441829681396
```

```

step: 50, loss: 0.2802959382534027
step: 60, loss: 0.24004794657230377
step: 70, loss: 0.37125399708747864
step: 80, loss: 0.36260420083999634
step: 90, loss: 0.23833706974983215
step: 100, loss: 0.26158368587493896
step: 110, loss: 0.2756504416465759
step: 120, loss: 0.1633516252040863
step: 130, loss: 0.04069983959197998
step: 140, loss: 0.21382558345794678
step: 150, loss: 0.25135719776153564
step: 160, loss: 0.11425969749689102
step: 170, loss: 0.2470792979001999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6840855106888362, f1=0.6940639269406393,
best_f1=0.6940639269406393
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19872689247131348
step: 10, loss: 0.20420396327972412
step: 20, loss: 0.10219303518533707
step: 30, loss: 0.16116441786289215
step: 40, loss: 0.12971587479114532
step: 50, loss: 0.04814177751541138
step: 60, loss: 0.23858699202537537
step: 70, loss: 0.06347134709358215
step: 80, loss: 0.07395129650831223
step: 90, loss: 0.10514435172080994
step: 100, loss: 0.07808026671409607
step: 110, loss: 0.12528015673160553
step: 120, loss: 0.09051764011383057
step: 130, loss: 0.11086882650852203
step: 140, loss: 0.31807848811149597
step: 150, loss: 0.1611425131559372
step: 160, loss: 0.2059711366891861
step: 170, loss: 0.10881996899843216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8055555555555555, f1=0.7662337662337663,
best_f1=0.7662337662337663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16851688921451569
step: 10, loss: 0.11338649690151215
step: 20, loss: 0.008791041560471058
step: 30, loss: 0.14452677965164185
step: 40, loss: 0.147104412317276
step: 50, loss: 0.07970519363880157
step: 60, loss: 0.17001500725746155
step: 70, loss: 0.1762019246816635
step: 80, loss: 0.032352514564991
step: 90, loss: 0.1265178918838501
step: 100, loss: 0.10732864588499069
step: 110, loss: 0.07663041353225708
step: 120, loss: 0.11481545865535736
step: 130, loss: 0.16858725249767303
step: 140, loss: 0.0877847746014595
step: 150, loss: 0.01678486168384552
step: 160, loss: 0.025171302258968353
step: 170, loss: 0.06757165491580963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8214285714285715, f1=0.7870967741935485,
best_f1=0.7870967741935485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01669536903500557
step: 10, loss: 0.05820135399699211
step: 20, loss: 0.038322437554597855
step: 30, loss: 0.04886588454246521
step: 40, loss: 0.001685825176537037
step: 50, loss: 0.0436592735350132
step: 60, loss: 0.08892016112804413
step: 70, loss: 0.02913646772503853
step: 80, loss: 0.11724204570055008
step: 90, loss: 0.0699867308139801
step: 100, loss: 0.021856151521205902

```



```

step: 110, loss: 0.10658933967351913
step: 120, loss: 0.007168088108301163
step: 130, loss: 0.049237463623285294
step: 140, loss: 0.04039390757679939
step: 150, loss: 0.2013874351978302
step: 160, loss: 0.3081703782081604
step: 170, loss: 0.029125496745109558
/usr/local/lib/python3.7/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=0.7990314769975787, f1=0.7981438515081206,
best_f1=0.7870967741935485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02525138109922409
step: 10, loss: 0.009125460870563984
step: 20, loss: 0.011141757480800152
step: 30, loss: 0.03320121765136719
step: 40, loss: 0.030279723927378654
step: 50, loss: 0.06303457170724869
step: 60, loss: 0.07214092463254929
step: 70, loss: 0.1356240212917328
step: 80, loss: 0.11028733104467392
step: 90, loss: 0.10518478602170944
step: 100, loss: 0.004558495711535215
step: 110, loss: 0.06845632195472717
step: 120, loss: 0.01653580367565155
step: 130, loss: 0.08092346042394638
step: 140, loss: 0.005918075796216726
step: 150, loss: 0.05408395826816559
step: 160, loss: 0.030507778748869896
step: 170, loss: 0.006138310767710209
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8400954653937948, f1=0.8349056603773586,
best_f1=0.8349056603773586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007116495165973902
step: 10, loss: 0.007245792541652918
step: 20, loss: 0.0011055454378947616

```

```

step: 30, loss: 0.07530611753463745
step: 40, loss: 0.04836051166057587
step: 50, loss: 0.12793928384780884
step: 60, loss: 0.05729758366942406
step: 70, loss: 0.08264121413230896
step: 80, loss: 0.015393896959722042
step: 90, loss: 0.0343746617436409
step: 100, loss: 0.03781495615839958
step: 110, loss: 0.000913378142286092
step: 120, loss: 0.011082606390118599
step: 130, loss: 0.03578127920627594
step: 140, loss: 0.01026011910289526
step: 150, loss: 0.009372921660542488
step: 160, loss: 0.07153519243001938
step: 170, loss: 0.005173951853066683
/usr/local/lib/python3.7/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=0.8264840182648402, f1=0.8103448275862069,
best_f1=0.8349056603773586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054857260547578335
step: 10, loss: 0.029697386547923088
step: 20, loss: 0.012381426058709621
step: 30, loss: 0.015387266874313354
step: 40, loss: 0.0031808128114789724
step: 50, loss: 0.0027837888337671757
step: 60, loss: 0.006545432843267918
step: 70, loss: 0.0015763939591124654
step: 80, loss: 0.014292526058852673
step: 90, loss: 0.0005340272327885032
step: 100, loss: 0.0003062500327359885
step: 110, loss: 0.00777302123606205
step: 120, loss: 0.007039709948003292
step: 130, loss: 0.14833760261535645
step: 140, loss: 0.0019779428839683533
step: 150, loss: 0.0035211616195738316
step: 160, loss: 0.0008057645754888654
step: 170, loss: 0.07185033708810806
/usr/local/lib/python3.7/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=0.8292682926829268, f1=0.832535885167464,
best_f1=0.8349056603773586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002790105761960149
step: 10, loss: 0.005218302831053734
step: 20, loss: 0.0009685410186648369
step: 30, loss: 0.007634647656232119
step: 40, loss: 0.0002717945317272097
step: 50, loss: 0.0030034708324819803
step: 60, loss: 0.0016901621129363775
step: 70, loss: 0.07803653180599213
step: 80, loss: 0.04133257642388344
step: 90, loss: 0.01611235737800598
step: 100, loss: 0.038357123732566833
step: 110, loss: 0.09154002368450165
step: 120, loss: 0.0019814905244857073
step: 130, loss: 0.0032446300610899925
step: 140, loss: 0.0012712335446849465
step: 150, loss: 0.018026217818260193
step: 160, loss: 0.07318489998579025
step: 170, loss: 0.003695968072861433
/usr/local/lib/python3.7/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=0.8329048843187661, f1=0.8159203980099503,
best_f1=0.8349056603773586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00044966614223085344
step: 10, loss: 0.05922802537679672
step: 20, loss: 0.006146234460175037
step: 30, loss: 0.004466914106160402
step: 40, loss: 0.3582083284854889
step: 50, loss: 0.0033427411690354347
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.11434117704629898
step: 70, loss: 0.005646837409585714
step: 80, loss: 0.03148077055811882
step: 90, loss: 0.16642989218235016
step: 100, loss: 0.04577501490712166
step: 110, loss: 0.002831944264471531
step: 120, loss: 0.005162048619240522
step: 130, loss: 0.06472277641296387
step: 140, loss: 0.027910472825169563
step: 150, loss: 0.003930514212697744
step: 160, loss: 0.04825440049171448
step: 170, loss: 0.0067564211785793304
/usr/local/lib/python3.7/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=0.8391959798994975, f1=0.8284313725490197,
best_f1=0.8349056603773586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08338035643100739
step: 10, loss: 0.0005374486208893359
step: 20, loss: 0.05324940383434296
step: 30, loss: 0.0173339881002903
step: 40, loss: 0.0005977389519102871
step: 50, loss: 0.1267784833908081
step: 60, loss: 0.007099096197634935
step: 70, loss: 0.030992936342954636
step: 80, loss: 0.007006456609815359
step: 90, loss: 0.01330125518143177
step: 100, loss: 0.0004099475627299398
step: 110, loss: 0.19020982086658478
step: 120, loss: 0.0021620430052280426
step: 130, loss: 0.015650395303964615
step: 140, loss: 0.15174271166324615
step: 150, loss: 0.004405399318784475
step: 160, loss: 0.00631765928119421
step: 170, loss: 0.02504870854318142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.85, f1=0.827250608272506, best_f1=0.827250608272506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027305150404572487
step: 10, loss: 0.004237083252519369
step: 20, loss: 0.016345929354429245
step: 30, loss: 0.0008354095625691116
step: 40, loss: 0.008147676475346088
step: 50, loss: 0.06259884685277939
step: 60, loss: 0.004643680062144995
step: 70, loss: 0.001989803509786725
step: 80, loss: 0.002926286542788148
step: 90, loss: 0.0011334101436659694
step: 100, loss: 0.0016003402415663004
step: 110, loss: 0.035126522183418274
step: 120, loss: 0.0007128210854716599

```

```

step: 130, loss: 0.00018891951185651124
step: 140, loss: 0.0076798624359071255
step: 150, loss: 0.0061439406126737595
step: 160, loss: 0.0025447956286370754
step: 170, loss: 0.001462931977584958
/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.8168316831683169,
best_f1=0.827250608272506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013378010131418705
step: 10, loss: 0.0002453148481436074
step: 20, loss: 0.0002449951134622097
step: 30, loss: 0.0006971896509639919
step: 40, loss: 0.014299864880740643
step: 50, loss: 0.00022110773716121912
step: 60, loss: 0.019584346562623978
step: 70, loss: 0.12909774482250214
step: 80, loss: 0.0002792864106595516
step: 90, loss: 0.0019813182298094034
step: 100, loss: 0.0015386141603812575
step: 110, loss: 0.001377699663862586
step: 120, loss: 0.001818773802369833
step: 130, loss: 0.06062174588441849
step: 140, loss: 0.007238270714879036
step: 150, loss: 0.0050431243143975735
step: 160, loss: 0.00037040209281258285
step: 170, loss: 0.00015750306192785501
/usr/local/lib/python3.7/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=0.8472906403940887, f1=0.8321513002364066,
best_f1=0.827250608272506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014307043515145779
step: 10, loss: 0.00023842194059398025
step: 20, loss: 0.0020591916982084513
step: 30, loss: 0.0007706991746090353
step: 40, loss: 0.004769362509250641
step: 50, loss: 0.0024841593112796545
step: 60, loss: 0.010614032857120037
step: 70, loss: 0.0003399935376364738
step: 80, loss: 0.007616463582962751

```

```

step: 90, loss: 0.0002977746189571917
step: 100, loss: 0.02132870815694332
step: 110, loss: 0.00010390218085376546
step: 120, loss: 0.0392439067363739
step: 130, loss: 0.008248244412243366
step: 140, loss: 0.000967244675848633
step: 150, loss: 0.0019912244752049446
step: 160, loss: 0.008268075063824654
step: 170, loss: 0.002487505553290248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8521303258145363, f1=0.819047619047619,
best_f1=0.819047619047619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003653263265732676
step: 10, loss: 0.0022288160398602486
step: 20, loss: 0.010290100239217281
step: 30, loss: 0.000363676663255319
step: 40, loss: 0.00015779081149958074
step: 50, loss: 0.23856402933597565
step: 60, loss: 0.0012345489813014865
step: 70, loss: 0.00244503072462976
step: 80, loss: 0.014253386296331882
step: 90, loss: 0.0002728019608184695
step: 100, loss: 0.000389839056879282
step: 110, loss: 0.0016581491800025105
step: 120, loss: 0.018471600487828255
step: 130, loss: 0.0025899969041347504
step: 140, loss: 0.00025658091180957854
step: 150, loss: 0.001151408301666379
step: 160, loss: 0.0009832491632550955
step: 170, loss: 0.023799624294042587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.8550000000000001, f1=0.8226950354609929,
best_f1=0.8226950354609929

```

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

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

step: 0, loss: 0.002358119236305356
step: 10, loss: 0.0017186697805300355
step: 20, loss: 0.04167560860514641
step: 30, loss: 0.005615185014903545
step: 40, loss: 0.0005504077416844666
step: 50, loss: 0.0007703137816861272
step: 60, loss: 0.02886911667883396
step: 70, loss: 0.0002012608601944521
step: 80, loss: 0.03665488213300705
step: 90, loss: 0.002948418725281954
step: 100, loss: 0.0010323271853849292
step: 110, loss: 0.0002132568188244477
step: 120, loss: 0.018061691895127296
step: 130, loss: 0.0018824625294655561
step: 140, loss: 0.0004470219719223678
step: 150, loss: 0.0001967744465218857
step: 160, loss: 0.000910138594917953
step: 170, loss: 0.0017684840131551027

/usr/local/lib/python3.7/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=0.8535980148883374, f1=0.822429906542056,
best_f1=0.8226950354609929

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, 258.46it/s]
load_f1 = 0.3824451410658307
real_f1 = 0.3753784056508577
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, 222.92it/s]

```

4.3 DITTO DIRTY

4.3.1 DBLP-ACM - Running the matcher

```

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

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

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

```



```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

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

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

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

step: 0, loss: 0.6056370735168457

step: 10, loss: 0.6293179988861084

step: 20, loss: 0.36575016379356384

```

step: 30, loss: 0.14814616739749908
step: 40, loss: 0.17232708632946014
step: 50, loss: 0.12908099591732025
step: 60, loss: 0.01190928928554058
step: 70, loss: 0.3077264130115509
step: 80, loss: 0.09899953007698059
step: 90, loss: 0.15782415866851807
step: 100, loss: 0.002218260895460844
step: 110, loss: 0.30982083082199097
step: 120, loss: 0.01583751104772091
step: 130, loss: 0.0074255564250051975
step: 140, loss: 0.002148691564798355
step: 150, loss: 0.012882286682724953
step: 160, loss: 0.032314762473106384
step: 170, loss: 0.10326699167490005
step: 180, loss: 0.018190741539001465
step: 190, loss: 0.08856332302093506
step: 200, loss: 0.03437814489006996
step: 210, loss: 0.015237783081829548
step: 220, loss: 0.006110935937613249
step: 230, loss: 0.06941917538642883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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=0.9743016759776536, f1=0.9742441209406495,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007266838103532791
step: 10, loss: 0.006769582629203796
step: 20, loss: 0.06343735009431839
step: 30, loss: 0.16399994492530823
step: 40, loss: 0.13637131452560425
step: 50, loss: 0.003564100479707122
step: 60, loss: 0.005283007863909006
step: 70, loss: 0.022467879578471184
step: 80, loss: 0.020344354212284088
step: 90, loss: 0.010250930674374104
step: 100, loss: 0.1304043084383011
step: 110, loss: 0.11160264909267426
step: 120, loss: 0.04009033739566803
step: 130, loss: 0.10339698940515518
step: 140, loss: 0.0060158888809382915

```

```

step: 150, loss: 0.054149385541677475
step: 160, loss: 0.09153623133897781
step: 170, loss: 0.003074592910706997
step: 180, loss: 0.020949196070432663
step: 190, loss: 0.003151176031678915
step: 200, loss: 0.002864327048882842
step: 210, loss: 0.001607992104254663
step: 220, loss: 0.11748319864273071
step: 230, loss: 0.03432760387659073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9786276715410572, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027331821620464325
step: 10, loss: 0.004326396156102419
step: 20, loss: 0.0017150931525975466
step: 30, loss: 0.022445840761065483
step: 40, loss: 0.149066761136055
step: 50, loss: 0.006686837878078222
step: 60, loss: 0.007098785601556301
step: 70, loss: 0.003358318470418453
step: 80, loss: 0.013181246817111969
step: 90, loss: 0.031475696712732315
step: 100, loss: 0.0008939526160247624
step: 110, loss: 0.0009209030540660024
step: 120, loss: 0.07997879385948181
step: 130, loss: 0.0003838458505924791
step: 140, loss: 0.13378943502902985
step: 150, loss: 0.015187425538897514
step: 160, loss: 0.02332584746181965
step: 170, loss: 0.018084973096847534
step: 180, loss: 0.004813734907656908
step: 190, loss: 0.0019759763963520527
step: 200, loss: 0.01592741161584854
step: 210, loss: 0.0949174091219902
step: 220, loss: 0.0012738234363496304
step: 230, loss: 0.0010439285542815924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, 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: 0.0004920278443023562
step: 10, loss: 0.0036508056800812483
step: 20, loss: 0.0038803115021437407
step: 30, loss: 0.0007683084695599973
step: 40, loss: 0.005148646887391806
step: 50, loss: 0.0015075235860422254
step: 60, loss: 0.005435311701148748
step: 70, loss: 0.0006458857678808272
step: 80, loss: 0.0003402431611903012
step: 90, loss: 0.00232272339053452
step: 100, loss: 0.000829377502668649
step: 110, loss: 0.0007550569716840982
step: 120, loss: 0.018913976848125458
step: 130, loss: 0.0025744158774614334
step: 140, loss: 0.009234361350536346
step: 150, loss: 0.12453585118055344
step: 160, loss: 0.048773393034935
step: 170, loss: 0.010929527692496777
step: 180, loss: 0.0018024002201855183
step: 190, loss: 0.0007295524701476097
step: 200, loss: 0.0009074269328266382
step: 210, loss: 0.06742016226053238
step: 220, loss: 0.00043674954213202
step: 230, loss: 0.016192186623811722
/usr/local/lib/python3.7/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=0.9810479375696767, f1=0.9788182831661093,
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: 0.002303107175976038
step: 10, loss: 0.011123333126306534
step: 20, loss: 0.00467123044654727
step: 30, loss: 0.0004058858612552285
step: 40, loss: 0.00106969999615103
step: 50, loss: 0.07257866859436035
step: 60, loss: 0.19717325270175934

```

```

step: 70, loss: 0.0637601688504219
step: 80, loss: 0.0015056064585223794
step: 90, loss: 0.004838668275624514
step: 100, loss: 0.00029304451891221106
step: 110, loss: 0.0003986556548625231
step: 120, loss: 0.00010138192737940699
step: 130, loss: 0.001707280520349741
step: 140, loss: 0.0012395920930430293
step: 150, loss: 0.003907155245542526
step: 160, loss: 0.0003495910787023604
step: 170, loss: 0.0007875073351897299
step: 180, loss: 0.0027524156030267477
step: 190, loss: 0.08966708928346634
step: 200, loss: 0.008698795922100544
step: 210, loss: 0.0017172598745673895
step: 220, loss: 0.0010231946362182498
step: 230, loss: 0.019473129883408546
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9775784753363228,
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: 0.015712551772594452
step: 10, loss: 0.0005740388296544552
step: 20, loss: 0.0008528043981641531
step: 30, loss: 0.00018961270689032972
step: 40, loss: 0.00017120172560680658
step: 50, loss: 0.003145252587273717
step: 60, loss: 0.00018507888307794929
step: 70, loss: 0.013669658452272415
step: 80, loss: 0.0002762811491265893
step: 90, loss: 0.0038824561052024364
step: 100, loss: 0.12815499305725098
step: 110, loss: 0.0021966188214719296
step: 120, loss: 0.002819980261847377
step: 130, loss: 0.0003247524728067219
step: 140, loss: 0.0002083141589537263
step: 150, loss: 0.010196898132562637
step: 160, loss: 0.00301332026720047
step: 170, loss: 0.004567326512187719
step: 180, loss: 0.006121943239122629
step: 190, loss: 0.008577240630984306
step: 200, loss: 0.00048111239448189735
step: 210, loss: 0.021230023354291916
step: 220, loss: 0.0005066648009233177

```

```

step: 230, loss: 0.10058071464300156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9808342728297633, f1=0.9831271091113611,
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: 0.006718407850712538
step: 10, loss: 9.771987970452756e-05
step: 20, loss: 0.0003196171310264617
step: 30, loss: 0.011462872847914696
step: 40, loss: 0.008944127708673477
step: 50, loss: 0.00023053355107549578
step: 60, loss: 9.693214815342799e-05
step: 70, loss: 0.04481683298945427
step: 80, loss: 0.007095706649124622
step: 90, loss: 0.0003226273984182626
step: 100, loss: 0.00022904275101609528
step: 110, loss: 0.002594518009573221
step: 120, loss: 8.328136027557775e-05
step: 130, loss: 0.0002205902710556984
step: 140, loss: 0.0003185905807185918
step: 150, loss: 0.008873892948031425
step: 160, loss: 0.044825974851846695
step: 170, loss: 0.0003548469103407115
step: 180, loss: 0.0024593521375209093
step: 190, loss: 0.00012908531061839312
step: 200, loss: 0.09209080785512924
step: 210, loss: 6.16105244262144e-05
step: 220, loss: 0.0009169374243356287
step: 230, loss: 0.00034096671151928604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853768278965129, f1=0.9831271091113611,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016253253852482885
step: 10, loss: 0.0001324293843936175
step: 20, loss: 0.00041961984243243933

```

```

step: 30, loss: 5.8799138059839606e-05
step: 40, loss: 0.001195419579744339
step: 50, loss: 0.0003247909771744162
step: 60, loss: 5.6272056099260226e-05
step: 70, loss: 0.00033310227445326746
step: 80, loss: 0.00016820442397147417
step: 90, loss: 4.679046105593443e-05
step: 100, loss: 0.00027122424216941
step: 110, loss: 0.0014944886788725853
step: 120, loss: 0.05562291294336319
step: 130, loss: 0.0015173153951764107
step: 140, loss: 7.207002636278048e-05
step: 150, loss: 9.652655717218295e-05
step: 160, loss: 0.00010690752242226154
step: 170, loss: 0.00023748379317112267
step: 180, loss: 0.03155509755015373
step: 190, loss: 0.006019368767738342
step: 200, loss: 0.00039650406688451767
step: 210, loss: 0.011317658238112926
step: 220, loss: 0.0002485725563019514
step: 230, loss: 0.00025390961673110723
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9796839729119639,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012246744881849736
step: 10, loss: 7.20412062946707e-05
step: 20, loss: 0.0019273774232715368
step: 30, loss: 7.426389493048191e-05
step: 40, loss: 0.0498676598072052
step: 50, loss: 5.619693547487259e-05
step: 60, loss: 5.4545900638913736e-05
step: 70, loss: 0.002563092391937971
step: 80, loss: 4.0142240322893485e-05
step: 90, loss: 6.953734555281699e-05
step: 100, loss: 0.0005077866371721029
step: 110, loss: 4.0994327719090506e-05
step: 120, loss: 0.0006289223674684763
step: 130, loss: 0.00011065021681133658
step: 140, loss: 0.00010760335135273635
step: 150, loss: 0.0016928586410358548
step: 160, loss: 0.00040406666812486947
step: 170, loss: 0.00022437766892835498
step: 180, loss: 0.02725670114159584

```

```

step: 190, loss: 0.00018264645768795162
step: 200, loss: 0.000987858627922833
step: 210, loss: 0.00011145542521262541
step: 220, loss: 5.425386916613206e-05
step: 230, loss: 0.00011999435810139403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887640449438202, f1=0.9787709497206705,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019774981774389744
step: 10, loss: 0.00012866769975516945
step: 20, loss: 0.0011339937336742878
step: 30, loss: 6.575524457730353e-05
step: 40, loss: 0.0001182197593152523
step: 50, loss: 0.0007097179186530411
step: 60, loss: 0.0001575362402945757
step: 70, loss: 0.0019310771021991968
step: 80, loss: 4.637044548871927e-05
step: 90, loss: 0.00034592190058901906
step: 100, loss: 4.655586599255912e-05
step: 110, loss: 0.0001249017077498138
step: 120, loss: 0.0002616362180560827
step: 130, loss: 5.5630884162383154e-05
step: 140, loss: 0.03235065937042236
step: 150, loss: 0.013913697563111782
step: 160, loss: 9.514049452263862e-05
step: 170, loss: 2.8821748855989426e-05
step: 180, loss: 7.450709381373599e-05
step: 190, loss: 0.01676896959543228
step: 200, loss: 3.198775084456429e-05
step: 210, loss: 0.0011539385886862874
step: 220, loss: 3.698311047628522e-05
step: 230, loss: 0.005149944685399532
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.977728285077951,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.936473750509322e-05
step: 10, loss: 0.0001138945299317129
step: 20, loss: 0.0033621834591031075
step: 30, loss: 0.0001371887483401224
step: 40, loss: 3.283745172666386e-05
step: 50, loss: 5.140231951372698e-05
step: 60, loss: 0.0007478748448193073
step: 70, loss: 0.0001389672397635877
step: 80, loss: 2.4150607714545913e-05
step: 90, loss: 0.00028666973230428994
step: 100, loss: 3.7627003621309996e-05
step: 110, loss: 0.0016763049643486738
step: 120, loss: 2.5443190679652616e-05
step: 130, loss: 3.126105002593249e-05
step: 140, loss: 2.3886117560323328e-05
step: 150, loss: 0.0219686608761549
step: 160, loss: 0.0001495561155024916
step: 170, loss: 0.011256893165409565
step: 180, loss: 0.0003720054228324443
step: 190, loss: 9.49113309616223e-05
step: 200, loss: 0.0003647645644377917
step: 210, loss: 2.5998233468271792e-05
step: 220, loss: 6.334439967758954e-05
step: 230, loss: 3.796183955273591e-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=0.9876265466816648, f1=0.9808773903262092,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.018943819683045e-05
step: 10, loss: 2.63782094407361e-05
step: 20, loss: 3.0453324143309146e-05
step: 30, loss: 6.694934563711286e-05
step: 40, loss: 9.092362597584724e-05
step: 50, loss: 0.0031416541896760464
step: 60, loss: 0.00012696215708274394
step: 70, loss: 2.703004065551795e-05
step: 80, loss: 0.00011809767602244392
step: 90, loss: 2.869513446057681e-05
step: 100, loss: 2.6780282496474683e-05
step: 110, loss: 2.9041539164609276e-05
step: 120, loss: 2.946248605439905e-05
step: 130, loss: 3.419690256123431e-05
step: 140, loss: 7.892473513493314e-05

```

```

step: 150, loss: 4.087861452717334e-05
step: 160, loss: 7.774239202262834e-05
step: 170, loss: 5.666236756951548e-05
step: 180, loss: 0.00010081313666887581
step: 190, loss: 4.415996227180585e-05
step: 200, loss: 2.0451461750781164e-05
step: 210, loss: 3.6740777431987226e-05
step: 220, loss: 4.1205570596503094e-05
step: 230, loss: 0.058129873126745224
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9776785714285714,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.327547013526782e-05
step: 10, loss: 0.0007161831599660218
step: 20, loss: 8.330623677466065e-05
step: 30, loss: 4.800777605851181e-05
step: 40, loss: 5.169233190827072e-05
step: 50, loss: 0.00021628252579830587
step: 60, loss: 0.0009602637146599591
step: 70, loss: 3.1450930691789836e-05
step: 80, loss: 0.00014591308718081564
step: 90, loss: 7.288562483154237e-05
step: 100, loss: 3.916150672012009e-05
step: 110, loss: 0.043661586940288544
step: 120, loss: 0.026690971106290817
step: 130, loss: 3.640561772044748e-05
step: 140, loss: 0.00032829135307110846
step: 150, loss: 2.654961463122163e-05
step: 160, loss: 0.0005928395548835397
step: 170, loss: 3.574613219825551e-05
step: 180, loss: 9.341479017166421e-05
step: 190, loss: 2.6802348656929098e-05
step: 200, loss: 0.00029762787744402885
step: 210, loss: 1.7624137399252504e-05
step: 220, loss: 0.0001082158341887407
step: 230, loss: 0.00017255364218726754
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9743016759776536,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.086192737100646e-05
step: 10, loss: 5.183641769690439e-05
step: 20, loss: 0.0001064466341631487
step: 30, loss: 0.0006210324936546385
step: 40, loss: 0.008710863068699837
step: 50, loss: 4.880769120063633e-05
step: 60, loss: 5.692314152838662e-05
step: 70, loss: 3.382021895959042e-05
step: 80, loss: 3.138455576845445e-05
step: 90, loss: 2.8318818294792436e-05
step: 100, loss: 3.910676241503097e-05
step: 110, loss: 0.001834893599152565
step: 120, loss: 1.646186865400523e-05
step: 130, loss: 2.3390673959511332e-05
step: 140, loss: 2.940662670880556e-05
step: 150, loss: 7.219319377327338e-05
step: 160, loss: 0.00649732630699873
step: 170, loss: 3.943030606023967e-05
step: 180, loss: 4.641664054361172e-05
step: 190, loss: 3.286647915956564e-05
step: 200, loss: 2.23996630666079e-05
step: 210, loss: 0.007171670440584421
step: 220, loss: 1.94978310901206e-05
step: 230, loss: 0.02157510630786419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9887640449438202, f1=0.9765363128491621,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011651076783891767
step: 10, loss: 1.4059120076126419e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 20, loss: 2.5163732061628252e-05
step: 30, loss: 3.298962110420689e-05
step: 40, loss: 3.458792343735695e-05
step: 50, loss: 0.04358305409550667
step: 60, loss: 3.5150045732734725e-05
step: 70, loss: 2.8255339202587493e-05
step: 80, loss: 0.0001155599020421505
step: 90, loss: 5.846595740877092e-05
step: 100, loss: 3.338849273859523e-05
step: 110, loss: 2.1732923414674588e-05
step: 120, loss: 0.00021006960014346987
step: 130, loss: 0.0022959711495786905

```

```

step: 140, loss: 3.2356019801227376e-05
step: 150, loss: 6.958210724405944e-05
step: 160, loss: 0.00011708229430951178
step: 170, loss: 1.9024781067855656e-05
step: 180, loss: 7.361547613982111e-05
step: 190, loss: 0.0011426678393036127
step: 200, loss: 3.502663457766175e-05
step: 210, loss: 0.0004389045643620193
step: 220, loss: 0.0006565917865373194
step: 230, loss: 2.0287530787754804e-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=0.9876265466816648, f1=0.9764837625979844,
best_f1=0.9787709497206705
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, 189.43it/s]
load_f1 = 0.9876543209876544
real_f1 = 0.9876543209876544

```

```
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 [00:19, 223.02it/s]
```

4.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

```
Downloading: 100% 433/433 [00:00<00:00, 396kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.53MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 68.0MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.
```

```

Defaults for this optimization level are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6145920753479004
step: 10, loss: 0.5114970803260803
step: 20, loss: 0.4955842196941376
step: 30, loss: 0.09067577123641968
step: 40, loss: 0.18826837837696075
step: 50, loss: 0.14789895713329315
step: 60, loss: 0.06626962125301361
step: 70, loss: 0.08287877589464188
step: 80, loss: 0.04762142151594162
step: 90, loss: 0.08797973394393921
step: 100, loss: 0.018702367320656776
step: 110, loss: 0.021284066140651703
step: 120, loss: 0.10423209518194199
step: 130, loss: 0.07367071509361267

```

```

step: 140, loss: 0.10319560766220093
step: 150, loss: 0.0398981086909771
step: 160, loss: 0.05601587891578674
step: 170, loss: 0.21834737062454224
step: 180, loss: 0.15650399029254913
step: 190, loss: 0.02347695082426071
step: 200, loss: 0.12969337403774261
step: 210, loss: 0.09391116350889206
step: 220, loss: 0.26244887709617615
step: 230, loss: 0.14446307718753815
step: 240, loss: 0.05160650610923767
step: 250, loss: 0.01743297278881073
step: 260, loss: 0.08647245168685913
step: 270, loss: 0.013230499811470509
step: 280, loss: 0.0333087220788002
step: 290, loss: 0.05635542422533035
step: 300, loss: 0.020965993404388428
step: 310, loss: 0.17244215309619904
step: 320, loss: 0.11697622388601303
step: 330, loss: 0.009970037266612053
step: 340, loss: 0.07498627156019211
step: 350, loss: 0.04707838594913483
step: 360, loss: 0.08101282268762589
step: 370, loss: 0.1595173180103302
step: 380, loss: 0.03268000856041908
step: 390, loss: 0.10689709335565567
step: 400, loss: 0.2673640847206116
step: 410, loss: 0.07610037922859192
step: 420, loss: 0.14710059762001038
step: 430, loss: 0.16564489901065826
step: 440, loss: 0.03441037982702255
step: 450, loss: 0.0036770962178707123
step: 460, loss: 0.05803699418902397
step: 470, loss: 0.0786772072315216
step: 480, loss: 0.03495178371667862
step: 490, loss: 0.15603376924991608
step: 500, loss: 0.08328594267368317
step: 510, loss: 0.06563229858875275
step: 520, loss: 0.023767346516251564
step: 530, loss: 0.00602251011878252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9423255813953489, f1=0.9417249417249418,

```

```
best_f1=0.9417249417249418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15640315413475037
step: 10, loss: 0.02647002600133419
step: 20, loss: 0.013672594912350178
step: 30, loss: 0.016396917402744293
step: 40, loss: 0.0482579804956913
step: 50, loss: 0.08824460953474045
step: 60, loss: 0.028206000104546547
step: 70, loss: 0.03358931466937065
step: 80, loss: 0.11721658706665039
step: 90, loss: 0.007796811405569315
step: 100, loss: 0.03986578434705734
step: 110, loss: 0.0395393930375576
step: 120, loss: 0.02521379105746746
step: 130, loss: 0.1570352017879486
step: 140, loss: 0.029445556923747063
step: 150, loss: 0.045546844601631165
step: 160, loss: 0.011713759042322636
step: 170, loss: 0.012347826734185219
step: 180, loss: 0.0151299349963665
step: 190, loss: 0.024709150195121765
step: 200, loss: 0.015582576394081116
step: 210, loss: 0.16685552895069122
step: 220, loss: 0.11113713681697845
step: 230, loss: 0.007935851812362671
step: 240, loss: 0.09750805050134659
step: 250, loss: 0.05341499671339989
step: 260, loss: 0.0012766573345288634
step: 270, loss: 0.03176480531692505
step: 280, loss: 0.010968275368213654
step: 290, loss: 0.005134630482643843
step: 300, loss: 0.04383376985788345
step: 310, loss: 0.007724924013018608
step: 320, loss: 0.12372826039791107
step: 330, loss: 0.08819888532161713
step: 340, loss: 0.032322414219379425
step: 350, loss: 0.0032471567392349243
step: 360, loss: 0.08904856443405151
step: 370, loss: 0.0744175910949707
step: 380, loss: 0.027375059202313423
step: 390, loss: 0.02540719136595726
step: 400, loss: 0.02252630703151226
step: 410, loss: 0.12243852764368057
step: 420, loss: 0.030447278171777725
step: 430, loss: 0.006653666030615568
```



```

step: 440, loss: 0.20687197148799896
step: 450, loss: 0.012024989351630211
step: 460, loss: 0.029530078172683716
step: 470, loss: 0.02878805808722973
step: 480, loss: 0.1547161191701889
step: 490, loss: 0.04956959933042526
step: 500, loss: 0.22905559837818146
step: 510, loss: 0.021428581327199936
step: 520, loss: 0.06646519154310226
step: 530, loss: 0.013676845468580723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9450045829514206, f1=0.9349704411095952,
best_f1=0.9349704411095952
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02544659934937954
step: 10, loss: 0.06770731508731842
step: 20, loss: 0.06114093214273453
step: 30, loss: 0.0363004244863987
step: 40, loss: 0.008076019585132599
step: 50, loss: 0.04942983016371727
step: 60, loss: 0.02096698433160782
step: 70, loss: 0.008202841505408287
step: 80, loss: 0.004795425105839968
step: 90, loss: 0.010576635599136353
step: 100, loss: 0.02045898512005806
step: 110, loss: 0.002563298912718892
step: 120, loss: 0.005375623237341642
step: 130, loss: 0.002877644030377269
step: 140, loss: 0.06997329741716385
step: 150, loss: 0.013141301460564137
step: 160, loss: 0.00981142371892929
step: 170, loss: 0.042110785841941833
step: 180, loss: 0.024719106033444405
step: 190, loss: 0.007107661105692387
step: 200, loss: 0.038404084742069244
step: 210, loss: 0.07562227547168732
step: 220, loss: 0.07764898240566254
step: 230, loss: 0.08371601998806
step: 240, loss: 0.01227688416838646
step: 250, loss: 0.04959537833929062

```

```

step: 260, loss: 0.013275256380438805
step: 270, loss: 0.011014130897819996
step: 280, loss: 0.1505504697561264
step: 290, loss: 0.0030419074464589357
step: 300, loss: 0.029711611568927765
step: 310, loss: 0.005876475479453802
step: 320, loss: 0.04202200844883919
step: 330, loss: 0.02713373862206936
step: 340, loss: 0.0045905387960374355
step: 350, loss: 0.0075764223001897335
step: 360, loss: 0.036212388426065445
step: 370, loss: 0.014925604686141014
step: 380, loss: 0.03263700008392334
step: 390, loss: 0.035242997109889984
step: 400, loss: 0.015719318762421608
step: 410, loss: 0.010945407673716545
step: 420, loss: 0.19833265244960785
step: 430, loss: 0.003940301481634378
step: 440, loss: 0.0023195818066596985
step: 450, loss: 0.0892622172832489
step: 460, loss: 0.035289961844682693
step: 470, loss: 0.17300580441951752
step: 480, loss: 0.001886469777673483
step: 490, loss: 0.0026651928201317787
step: 500, loss: 0.02286924421787262
step: 510, loss: 0.005991511512547731
step: 520, loss: 0.13980238139629364
step: 530, loss: 0.023901918902993202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9481961147086033, f1=0.9362292051756008,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027502281591296196
step: 10, loss: 0.0037151470314711332
step: 20, loss: 0.011156030930578709
step: 30, loss: 0.0013392178807407618
step: 40, loss: 0.03768511861562729
step: 50, loss: 0.0015733785694465041
step: 60, loss: 0.006371563300490379
step: 70, loss: 0.006387566216289997

```

step: 80, loss: 0.006065207067877054
step: 90, loss: 0.02255960740149021
step: 100, loss: 0.02236853539943695
step: 110, loss: 0.029784470796585083
step: 120, loss: 0.006127167493104935
step: 130, loss: 0.00311979535035789
step: 140, loss: 0.0013827773509547114
step: 150, loss: 0.0525667704641819
step: 160, loss: 0.06260811537504196
step: 170, loss: 0.009776576422154903
step: 180, loss: 0.006925389636307955
step: 190, loss: 0.12952393293380737
step: 200, loss: 0.017913205549120903
step: 210, loss: 0.013229429721832275
step: 220, loss: 0.007185324560850859
step: 230, loss: 0.08044709265232086
step: 240, loss: 0.0015363963320851326
step: 250, loss: 0.041155748069286346
step: 260, loss: 0.013259579427540302
step: 270, loss: 0.00954575464129448
step: 280, loss: 0.023662911728024483
step: 290, loss: 0.027593301609158516
step: 300, loss: 0.0001171483745565638
step: 310, loss: 0.0047628008760511875
step: 320, loss: 0.08198503404855728
step: 330, loss: 0.014716843143105507
step: 340, loss: 0.004795833025127649
step: 350, loss: 0.08957027643918991
step: 360, loss: 0.023539355024695396
step: 370, loss: 0.004931177943944931
step: 380, loss: 0.0033502262085676193
step: 390, loss: 0.0009761127294041216
step: 400, loss: 0.002545579569414258
step: 410, loss: 0.007046564947813749
step: 420, loss: 0.001453037024475634
step: 430, loss: 0.1628064066171646
step: 440, loss: 0.005500433035194874
step: 450, loss: 0.01893686130642891
step: 460, loss: 0.003718265099450946
step: 470, loss: 0.04445303976535797
step: 480, loss: 0.2264351099729538
step: 490, loss: 0.0052770269103348255
step: 500, loss: 0.003593944013118744
step: 510, loss: 0.03457548841834068
step: 520, loss: 0.08469858765602112
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 530, loss: 0.029390236362814903
/usr/local/lib/python3.7/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=0.9478060046189377, f1=0.9396591432519575,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020888440310955048
step: 10, loss: 0.1301560252904892
step: 20, loss: 0.005012364126741886
step: 30, loss: 0.005619258154183626
step: 40, loss: 0.012469911016523838
step: 50, loss: 0.0012130747782066464
step: 60, loss: 0.02928241528570652
step: 70, loss: 0.0035045607946813107
step: 80, loss: 0.0012225990649312735
step: 90, loss: 0.002162795513868332
step: 100, loss: 0.0019857161678373814
step: 110, loss: 0.00014084490248933434
step: 120, loss: 0.0024357885122299194
step: 130, loss: 0.00024234184820670635
step: 140, loss: 0.002785891992971301
step: 150, loss: 0.09816057235002518
step: 160, loss: 0.0012466927291825414
step: 170, loss: 0.01506613940000534
step: 180, loss: 0.0009893238311633468
step: 190, loss: 0.0005696074804291129
step: 200, loss: 0.007981283590197563
step: 210, loss: 0.002626928733661771
step: 220, loss: 0.0062655978836119175
step: 230, loss: 0.023187153041362762
step: 240, loss: 0.002004630398005247
step: 250, loss: 0.004352543968707323
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 260, loss: 0.00039429112803190947
step: 270, loss: 0.0023417265620082617
step: 280, loss: 0.029944950714707375
step: 290, loss: 0.12373227626085281
step: 300, loss: 0.0214021485298872
step: 310, loss: 0.008920508436858654
step: 320, loss: 0.11462520807981491
step: 330, loss: 0.03525318205356598
step: 340, loss: 0.006154017057269812
step: 350, loss: 0.004608019720762968
step: 360, loss: 0.0010339532746002078
step: 370, loss: 0.03706176578998566
step: 380, loss: 0.00036350652226246893
step: 390, loss: 0.00017165277677122504
```

```

step: 400, loss: 0.010529348626732826
step: 410, loss: 0.0024083047173917294
step: 420, loss: 0.01350437942892313
step: 430, loss: 0.004426778759807348
step: 440, loss: 0.008779894560575485
step: 450, loss: 0.002172653330489993
step: 460, loss: 0.00024393803323619068
step: 470, loss: 0.007876263000071049
step: 480, loss: 0.11996226012706757
step: 490, loss: 0.0003227478009648621
step: 500, loss: 0.05322647467255592
step: 510, loss: 0.0007050717831589282
step: 520, loss: 0.008946389891207218
step: 530, loss: 0.0017867913702502847
/usr/local/lib/python3.7/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=0.942095588235294, f1=0.9244851258581235,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019193051382899284
step: 10, loss: 0.000423215125920251
step: 20, loss: 0.04847110062837601
step: 30, loss: 0.00040082994382828474
step: 40, loss: 0.18270163238048553
step: 50, loss: 0.014922775328159332
step: 60, loss: 0.0007479963242076337
step: 70, loss: 0.004952699411660433
step: 80, loss: 0.09799090027809143
step: 90, loss: 0.004490083549171686
step: 100, loss: 0.05413204804062843
step: 110, loss: 0.04092201590538025
step: 120, loss: 0.001968792174011469
step: 130, loss: 0.02044561877846718
step: 140, loss: 0.0524638369679451
step: 150, loss: 0.0003560655750334263
step: 160, loss: 0.00016604231495875865
step: 170, loss: 0.030321499332785606
step: 180, loss: 0.022446246817708015
step: 190, loss: 0.008134210482239723
step: 200, loss: 0.0023560686968266964
step: 210, loss: 0.011343627236783504
step: 220, loss: 0.012868214398622513
step: 230, loss: 0.00043066678335890174
step: 240, loss: 0.06244657188653946
step: 250, loss: 0.00133999134413898

```

```

step: 260, loss: 0.0007505477988161147
step: 270, loss: 0.06225084885954857
step: 280, loss: 0.07670234143733978
step: 290, loss: 0.0006019973661750555
step: 300, loss: 0.003028338076546788
step: 310, loss: 0.0033909042831510305
step: 320, loss: 0.002184320241212845
step: 330, loss: 0.004274426493793726
step: 340, loss: 0.0976334661245346
step: 350, loss: 0.0003435480175539851
step: 360, loss: 0.003949319012463093
step: 370, loss: 0.0322648324072361
step: 380, loss: 0.0009606530074961483
step: 390, loss: 0.046073079109191895
step: 400, loss: 0.0001133174664573744
step: 410, loss: 0.0066047655418515205
step: 420, loss: 0.02405090630054474
step: 430, loss: 0.0025368130300194025
step: 440, loss: 0.0013293006923049688
step: 450, loss: 0.0007582625839859247
step: 460, loss: 0.00018495779659133404
step: 470, loss: 0.01069552917033434
step: 480, loss: 0.014161661267280579
step: 490, loss: 0.009948527440428734
step: 500, loss: 0.0007081009680405259
step: 510, loss: 0.012834623456001282
step: 520, loss: 0.010115033946931362
step: 530, loss: 0.0038545792922377586
/usr/local/lib/python3.7/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=0.9444444444444445, f1=0.9348729792147806,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027287373319268227
step: 10, loss: 0.0006954956916160882
step: 20, loss: 0.0033676535822451115
step: 30, loss: 0.0006912574172019958
step: 40, loss: 0.11959753930568695
step: 50, loss: 0.007553664036095142
step: 60, loss: 0.06302458047866821
step: 70, loss: 0.0018035239772871137
step: 80, loss: 0.010622917674481869
step: 90, loss: 0.006295275408774614
step: 100, loss: 5.338493792805821e-05
step: 110, loss: 0.003060809802263975

```

```
step: 120, loss: 0.000251447461778298
step: 130, loss: 0.0004952181479893625
step: 140, loss: 0.0005356041365303099
step: 150, loss: 0.0014402633532881737
step: 160, loss: 8.732249989407137e-05
step: 170, loss: 0.0006846781470812857
step: 180, loss: 0.0008567323093302548
step: 190, loss: 0.0005411678575910628
step: 200, loss: 8.349737618118525e-05
step: 210, loss: 0.007203255780041218
step: 220, loss: 0.13272036612033844
step: 230, loss: 0.038134872913360596
step: 240, loss: 0.005871471483260393
step: 250, loss: 0.007405213546007872
step: 260, loss: 0.006561345420777798
step: 270, loss: 0.015932787209749222
step: 280, loss: 0.02320544794201851
step: 290, loss: 0.035810697823762894
step: 300, loss: 0.0010094065219163895
step: 310, loss: 0.0007712900405749679
step: 320, loss: 0.011277472600340843
step: 330, loss: 0.012872875668108463
step: 340, loss: 0.02399352192878723
step: 350, loss: 0.001111515681259334
step: 360, loss: 0.007112578488886356
step: 370, loss: 0.001084078918211162
step: 380, loss: 0.010780488140881062
step: 390, loss: 0.0004626363515853882
step: 400, loss: 0.008736525662243366
step: 410, loss: 0.0024310068693012
step: 420, loss: 0.0001502220839029178
step: 430, loss: 0.0023483301047235727
step: 440, loss: 0.017221754416823387
step: 450, loss: 0.00041846130625344813
step: 460, loss: 0.0004774107364937663
step: 470, loss: 0.22970931231975555
step: 480, loss: 0.07772494852542877
step: 490, loss: 0.005121293012052774
step: 500, loss: 0.0098673515021801
step: 510, loss: 0.011409714818000793
step: 520, loss: 0.005022442899644375
step: 530, loss: 0.005184534005820751
/usr/local/lib/python3.7/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=0.9466911764705882, f1=0.9330296127562643,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 0.121160589158535
step: 10, loss: 0.005208613816648722
step: 20, loss: 0.00045690653496421874
step: 30, loss: 0.0007983780815266073
step: 40, loss: 0.043718043714761734
step: 50, loss: 0.0658392384648323
step: 60, loss: 0.000685787876136601
step: 70, loss: 0.0017725670477375388
step: 80, loss: 0.009057257324457169
step: 90, loss: 0.0030838288366794586
step: 100, loss: 0.0005241035250946879
step: 110, loss: 0.0007215893128886819
step: 120, loss: 0.011307785287499428
step: 130, loss: 0.0011028082808479667
step: 140, loss: 0.004659739322960377
step: 150, loss: 0.016086449846625328
step: 160, loss: 0.02073766104876995
step: 170, loss: 0.04063522070646286
step: 180, loss: 0.0015896253753453493
step: 190, loss: 0.0028493357822299004
step: 200, loss: 0.0029512662440538406
step: 210, loss: 0.004189124330878258
step: 220, loss: 0.0076401084661483765
step: 230, loss: 0.004102739505469799
step: 240, loss: 0.0005558698903769255
step: 250, loss: 0.00021349637245293707
step: 260, loss: 0.020317556336522102
step: 270, loss: 0.002883754437789321
step: 280, loss: 0.09574811905622482
step: 290, loss: 6.665159889962524e-05
step: 300, loss: 0.00029036711202934384
step: 310, loss: 0.0018807189771905541
step: 320, loss: 0.0007167813018895686
step: 330, loss: 0.003999779466539621
step: 340, loss: 6.076943100197241e-05
step: 350, loss: 0.10718249529600143
step: 360, loss: 0.0017082224367186427
step: 370, loss: 0.0011356030590832233
step: 380, loss: 0.005884664598852396
step: 390, loss: 0.14669418334960938
step: 400, loss: 0.0038521783426404
step: 410, loss: 0.0003501072060316801
step: 420, loss: 0.013468710705637932
step: 430, loss: 0.015796635299921036
step: 440, loss: 0.0018735987832769752
step: 450, loss: 0.0007046034443192184
```



```
step: 460, loss: 0.0005282313213683665
step: 470, loss: 0.000786761986091733
step: 480, loss: 0.000648730609100312
step: 490, loss: 0.006383582018315792
step: 500, loss: 0.0005312658613547683
step: 510, loss: 0.002637198893353343
step: 520, loss: 0.002286977833136916
step: 530, loss: 0.00039917120011523366
/usr/local/lib/python3.7/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=0.9417792268281323, f1=0.9347623485554521,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002178974449634552
step: 10, loss: 0.0007501555955968797
step: 20, loss: 0.002125277416780591
step: 30, loss: 0.00017249461961910129
step: 40, loss: 3.721309622051194e-05
step: 50, loss: 8.664221240906045e-05
step: 60, loss: 4.5516611862694845e-05
step: 70, loss: 0.0007515048491768539
step: 80, loss: 0.015372000634670258
step: 90, loss: 3.497496072668582e-05
step: 100, loss: 0.0005051757325418293
step: 110, loss: 0.00014855457993689924
step: 120, loss: 6.240960647119209e-05
step: 130, loss: 6.227532139746472e-05
step: 140, loss: 0.00029549331520684063
step: 150, loss: 0.00030626001534983516
step: 160, loss: 0.0002646410430315882
step: 170, loss: 0.005248101893812418
step: 180, loss: 4.354953489382751e-05
step: 190, loss: 0.00017067593580577523
step: 200, loss: 0.0005991065409034491
step: 210, loss: 5.842565224156715e-05
step: 220, loss: 0.000276994367595762
step: 230, loss: 0.009249777533113956
step: 240, loss: 6.946920620976016e-05
step: 250, loss: 0.0001057663030223921
step: 260, loss: 0.05984831973910332
step: 270, loss: 8.965528104454279e-05
step: 280, loss: 0.037818435579538345
step: 290, loss: 0.02929808758199215
step: 300, loss: 4.242160139256157e-05
step: 310, loss: 0.00450296513736248
```

```

step: 320, loss: 0.0007791363750584424
step: 330, loss: 0.0008677329169586301
step: 340, loss: 0.0008606127230450511
step: 350, loss: 0.024349097162485123
step: 360, loss: 0.0018771829782053828
step: 370, loss: 0.0029128468595445156
step: 380, loss: 0.00010260588896926492
step: 390, loss: 0.00012020440772175789
step: 400, loss: 0.012914231047034264
step: 410, loss: 0.00012681953376159072
step: 420, loss: 0.000103371414297726
step: 430, loss: 0.00018807625747285783
step: 440, loss: 0.0020336576271802187
step: 450, loss: 0.0003309926833026111
step: 460, loss: 0.000340341153787449
step: 470, loss: 0.00017589707567822188
step: 480, loss: 6.311624019872397e-05
step: 490, loss: 0.0024359161034226418
step: 500, loss: 0.0011746217496693134
step: 510, loss: 0.0005011546891182661
step: 520, loss: 0.0001511929731350392
step: 530, loss: 0.008394675329327583
/usr/local/lib/python3.7/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=0.9440145653163405, f1=0.9330316742081448,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019160262309014797
step: 10, loss: 0.03985920920968056
step: 20, loss: 3.0129394872346893e-05
step: 30, loss: 0.004027320072054863
step: 40, loss: 0.0022593841422349215
step: 50, loss: 5.115313615533523e-05
step: 60, loss: 0.009350724518299103
step: 70, loss: 3.276971619925462e-05
step: 80, loss: 4.692612856160849e-05
step: 90, loss: 0.000463175616459921
step: 100, loss: 3.764959183172323e-05
step: 110, loss: 5.587739360635169e-05
step: 120, loss: 0.0005292354617267847
step: 130, loss: 0.019514000043272972
step: 140, loss: 0.0061705633997917175
step: 150, loss: 0.0020420015789568424
step: 160, loss: 0.00014613018720410764
step: 170, loss: 0.00016611801402177662

```

```

step: 180, loss: 4.4410739064915106e-05
step: 190, loss: 3.780197585001588e-05
step: 200, loss: 0.010797337628901005
step: 210, loss: 6.933079566806555e-05
step: 220, loss: 0.05205285921692848
step: 230, loss: 0.0005049759056419134
step: 240, loss: 0.010681813582777977
step: 250, loss: 2.7376301659387536e-05
step: 260, loss: 0.0017134150257334113
step: 270, loss: 0.003409113734960556
step: 280, loss: 0.00030017216340638697
step: 290, loss: 0.000160520663484931
step: 300, loss: 2.9805400117766112e-05
step: 310, loss: 0.0006096602883189917
step: 320, loss: 0.0014381015207618475
step: 330, loss: 0.05545781925320625
step: 340, loss: 0.0032310534734278917
step: 350, loss: 0.023956572636961937
step: 360, loss: 0.0004121583478990942
step: 370, loss: 0.012148167937994003
step: 380, loss: 0.019493378698825836
step: 390, loss: 0.001657164772041142
step: 400, loss: 0.001167403650470078
step: 410, loss: 0.002547413809224963
step: 420, loss: 0.0005154449609108269
step: 430, loss: 0.0003521887701936066
step: 440, loss: 0.00345217389985919
step: 450, loss: 6.828410550951958e-05
step: 460, loss: 5.853126640431583e-05
step: 470, loss: 0.008678372018039227
step: 480, loss: 0.0009295828058384359
step: 490, loss: 0.0011029407614842057
step: 500, loss: 0.05833686888217926
step: 510, loss: 3.618426853790879e-05
step: 520, loss: 0.00010222114360658452
step: 530, loss: 0.00029197928961366415
/usr/local/lib/python3.7/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=0.9446529080675422, f1=0.9396914446002805,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039270843262784183
step: 10, loss: 0.08515433222055435
step: 20, loss: 0.0002876677899621427
step: 30, loss: 0.0017891019815579057

```

step: 40, loss: 0.000369192857760936
step: 50, loss: 0.00030452347709797323
step: 60, loss: 0.03689834848046303
step: 70, loss: 4.922880179947242e-05
step: 80, loss: 0.018627382814884186
step: 90, loss: 3.069877129746601e-05
step: 100, loss: 0.0019391275709494948
step: 110, loss: 0.0001409261894878
step: 120, loss: 0.000634649011772126
step: 130, loss: 0.00014986075984779745
step: 140, loss: 0.011362520046532154
step: 150, loss: 3.167104296153411e-05
step: 160, loss: 0.027868399396538734
step: 170, loss: 6.17465702816844e-05
step: 180, loss: 0.0009405092569068074
step: 190, loss: 0.003776292782276869
step: 200, loss: 0.02690928801894188
step: 210, loss: 0.0009915572591125965
step: 220, loss: 0.003485624212771654
step: 230, loss: 9.514558769296855e-05
step: 240, loss: 0.00011121352144982666
step: 250, loss: 1.9374951079953462e-05
step: 260, loss: 2.609055991342757e-05
step: 270, loss: 0.0024164330679923296
step: 280, loss: 0.00015454496315214783
step: 290, loss: 0.0010020805057138205
step: 300, loss: 0.0007992798346094787
step: 310, loss: 0.0008199394796974957
step: 320, loss: 1.9050856280955486e-05
step: 330, loss: 8.65229667397216e-05
step: 340, loss: 0.008910216391086578
step: 350, loss: 0.0002417811338091269
step: 360, loss: 2.9502605684683658e-05
step: 370, loss: 0.0003621435316745192
step: 380, loss: 1.649901423661504e-05
step: 390, loss: 5.798539496026933e-05
step: 400, loss: 0.004167822189629078
step: 410, loss: 0.00019622259424068034
step: 420, loss: 0.00014741467020940036
step: 430, loss: 8.956061356002465e-05
step: 440, loss: 0.0013185767456889153
step: 450, loss: 9.399418922839686e-05
step: 460, loss: 0.0073603251948952675
step: 470, loss: 2.2853922928334214e-05
step: 480, loss: 0.0031903348863124847
step: 490, loss: 0.006090480834245682
step: 500, loss: 0.0006076873978599906
step: 510, loss: 0.004333446733653545

```

step: 520, loss: 0.0017101895064115524
step: 530, loss: 2.375903022766579e-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=0.9418334108887855, f1=0.937528921795465,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.746479003690183e-05
step: 10, loss: 2.975069401145447e-05
step: 20, loss: 8.070170588325709e-05
step: 30, loss: 0.007953333668410778
step: 40, loss: 0.0005128878401592374
step: 50, loss: 0.030775975435972214
step: 60, loss: 0.004182418808341026
step: 70, loss: 0.015254628844559193
step: 80, loss: 4.048152550240047e-05
step: 90, loss: 2.8179556466056965e-05
step: 100, loss: 0.03129296749830246
step: 110, loss: 0.09015518426895142
step: 120, loss: 0.0002803835377562791
step: 130, loss: 0.0060492753982543945
step: 140, loss: 2.41240795730846e-05
step: 150, loss: 0.00010280551214236766
step: 160, loss: 5.1237882871646434e-05
step: 170, loss: 3.649702557595447e-05
step: 180, loss: 0.00013389850209932774
step: 190, loss: 6.047740316716954e-05
step: 200, loss: 0.0001213130381074734
step: 210, loss: 6.341766857076436e-05
step: 220, loss: 0.1948615163564682
step: 230, loss: 5.50870354345534e-05
step: 240, loss: 0.00019942515064030886
step: 250, loss: 6.183755613164976e-05
step: 260, loss: 5.643947224598378e-05
step: 270, loss: 2.526783646317199e-05
step: 280, loss: 0.00022600704687647521
step: 290, loss: 0.05834482982754707
step: 300, loss: 0.0007964209071360528
step: 310, loss: 0.000267757655819878
step: 320, loss: 4.800432361662388e-05
step: 330, loss: 0.0006495339912362397
step: 340, loss: 0.0008688624948263168
step: 350, loss: 0.0008283298229798675
step: 360, loss: 0.0014060883549973369
step: 370, loss: 0.003723327536135912

```

```
step: 380, loss: 0.00023588843760080636
step: 390, loss: 0.004642090760171413
step: 400, loss: 0.001736174221150577
step: 410, loss: 0.0002825424016918987
step: 420, loss: 0.009202118963003159
step: 430, loss: 0.013965564779937267
step: 440, loss: 0.005809273105114698
step: 450, loss: 0.00044474986498244107
step: 460, loss: 6.082710388000123e-05
step: 470, loss: 3.7213205359876156e-05
step: 480, loss: 0.16537131369113922
step: 490, loss: 0.0001765760243870318
step: 500, loss: 0.0008354299934580922
step: 510, loss: 0.008513922803103924
step: 520, loss: 2.282420427945908e-05
step: 530, loss: 1.4360808563651517e-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=0.943327239488117, f1=0.9347826086956521,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00571275083348155
step: 10, loss: 4.534816252999008e-05
step: 20, loss: 0.0003018188872374594
step: 30, loss: 6.319931708276272e-05
step: 40, loss: 0.0004420220502652228
step: 50, loss: 0.0032436223700642586
step: 60, loss: 0.00034562303335405886
step: 70, loss: 3.867493069265038e-05
step: 80, loss: 0.00016021441842895
step: 90, loss: 0.03556443005800247
step: 100, loss: 0.0005497168749570847
step: 110, loss: 0.0002721132477745414
step: 120, loss: 0.00012130582763347775
step: 130, loss: 3.960078174713999e-05
step: 140, loss: 0.0006186281680129468
step: 150, loss: 0.030012717470526695
step: 160, loss: 2.8757423933711834e-05
step: 170, loss: 0.00012781859550159425
step: 180, loss: 0.00013258485705591738
step: 190, loss: 5.818726640427485e-05
step: 200, loss: 9.90933840512298e-05
step: 210, loss: 0.0006137332529760897
step: 220, loss: 5.1142076699761674e-05
step: 230, loss: 0.0005616458947770298
```

```

step: 240, loss: 0.000377807387849316
step: 250, loss: 0.00047202332643792033
step: 260, loss: 0.0011832279851660132
step: 270, loss: 1.9281742424936965e-05
step: 280, loss: 0.00010366681817686185
step: 290, loss: 9.121531184064224e-05
step: 300, loss: 3.5773744457401335e-05
step: 310, loss: 0.0007156957290135324
step: 320, loss: 0.006881745997816324
step: 330, loss: 0.1325252652168274
step: 340, loss: 0.0017797357868403196
step: 350, loss: 0.00021303980611264706
step: 360, loss: 2.908956230385229e-05
step: 370, loss: 2.7055080863647163e-05
step: 380, loss: 3.208390626241453e-05
step: 390, loss: 0.0015618772013112903
step: 400, loss: 0.00039274070877581835
step: 410, loss: 0.0005443188128992915
step: 420, loss: 2.180344563385006e-05
step: 430, loss: 0.042719315737485886
step: 440, loss: 0.002888723975047469
step: 450, loss: 0.0007852959679439664
step: 460, loss: 0.0008050671312958002
step: 470, loss: 0.0001127851428464055
step: 480, loss: 4.665489905164577e-05
step: 490, loss: 4.1092389437835664e-05
step: 500, loss: 7.747601921437308e-05
step: 510, loss: 7.145893323468044e-05
step: 520, loss: 0.005018600728362799
step: 530, loss: 2.353903437324334e-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=0.9450549450549451, f1=0.9370185772541911,
best_f1=0.9362292051756008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002085028972942382
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 2.7401587431086227e-05
step: 20, loss: 2.533808583393693e-05
step: 30, loss: 2.5889963580993935e-05
step: 40, loss: 0.0035576599184423685
step: 50, loss: 0.001717897830531001
step: 60, loss: 1.3567386304202955e-05
step: 70, loss: 8.194782276405022e-05
step: 80, loss: 8.919946412788704e-05

```

```
step: 90, loss: 0.00028425382333807647
step: 100, loss: 0.00014447758439928293
step: 110, loss: 2.138272611773573e-05
step: 120, loss: 0.0030518544372171164
step: 130, loss: 0.0031750814523547888
step: 140, loss: 6.302983092609793e-05
step: 150, loss: 3.441411172389053e-05
step: 160, loss: 0.00029807808459736407
step: 170, loss: 4.485165482037701e-05
step: 180, loss: 2.246906115033198e-05
step: 190, loss: 0.00012647663243114948
step: 200, loss: 2.0265117200324312e-05
step: 210, loss: 3.093591294600628e-05
step: 220, loss: 2.1147889128769748e-05
step: 230, loss: 0.00019920161867048591
step: 240, loss: 0.0001926185068441555
step: 250, loss: 3.796525197685696e-05
step: 260, loss: 0.00017678721633274108
step: 270, loss: 0.00035949074663221836
step: 280, loss: 2.0563069483614527e-05
step: 290, loss: 5.4866482969373465e-05
step: 300, loss: 0.0002552731893956661
step: 310, loss: 0.015172204002737999
step: 320, loss: 0.00011086584709119052
step: 330, loss: 8.204748155549169e-05
step: 340, loss: 3.7944097130093724e-05
step: 350, loss: 8.504361176164821e-05
step: 360, loss: 0.03414090722799301
step: 370, loss: 0.0004282362060621381
step: 380, loss: 0.00030063246958889067
step: 390, loss: 0.0042626867070794106
step: 400, loss: 0.00024137651780620217
step: 410, loss: 3.127732270513661e-05
step: 420, loss: 0.0002606046327855438
step: 430, loss: 0.0009456119150854647
step: 440, loss: 0.021595366299152374
step: 450, loss: 8.285876538138837e-05
step: 460, loss: 8.224081830121577e-05
step: 470, loss: 2.7301182853989303e-05
step: 480, loss: 0.00013707716425415128
step: 490, loss: 1.9650562535389327e-05
step: 500, loss: 2.2663805793854408e-05
step: 510, loss: 0.0004012261051684618
step: 520, loss: 3.988224852946587e-05
step: 530, loss: 7.53677450120449e-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=0.9429484253765406, f1=0.9370185772541911,  
best_f1=0.9362292051756008  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 4.59910552308429e-05  
step: 10, loss: 2.1725149053963833e-05  
step: 20, loss: 0.00018370604084338993  
step: 30, loss: 0.00027406064327806234  
step: 40, loss: 0.019081465899944305  
step: 50, loss: 0.0019642726983875036  
step: 60, loss: 9.74502763710916e-05  
step: 70, loss: 0.0013561621308326721  
step: 80, loss: 0.0001250754576176405  
step: 90, loss: 2.8802403903682716e-05  
step: 100, loss: 3.633799497038126e-05  
step: 110, loss: 0.00017662155732978135  
step: 120, loss: 0.019913382828235626  
step: 130, loss: 0.09549637138843536  
step: 140, loss: 2.9130040275049396e-05  
step: 150, loss: 3.665430631372146e-05  
step: 160, loss: 5.0280115829082206e-05  
step: 170, loss: 4.3549356632865965e-05  
step: 180, loss: 4.813457780983299e-05  
step: 190, loss: 3.229264257242903e-05  
step: 200, loss: 0.00015886608161963522  
step: 210, loss: 0.0021119248121976852  
step: 220, loss: 0.0002632063115015626  
step: 230, loss: 0.0010826698271557689  
step: 240, loss: 3.860257857013494e-05  
step: 250, loss: 8.978343976195902e-05  
step: 260, loss: 0.006491706240922213  
step: 270, loss: 3.197781552444212e-05  
step: 280, loss: 0.0008097469108179212  
step: 290, loss: 1.646558303036727e-05  
step: 300, loss: 3.201375511707738e-05  
step: 310, loss: 0.00018475567048881203  
step: 320, loss: 0.00032959325471892953  
step: 330, loss: 4.877314859186299e-05  
step: 340, loss: 3.0132634492474608e-05  
step: 350, loss: 5.225426139077172e-05  
step: 360, loss: 0.006912435404956341  
step: 370, loss: 5.752022116212174e-05  
step: 380, loss: 0.00013125348777975887  
step: 390, loss: 0.016986722126603127  
step: 400, loss: 0.001587389619089663  
step: 410, loss: 0.0001599519164301455  
step: 420, loss: 8.327524119522423e-05
```

```

step: 430, loss: 0.00011675943824229762
step: 440, loss: 0.004311677068471909
step: 450, loss: 3.516333890729584e-05
step: 460, loss: 1.2595069165399764e-05
step: 470, loss: 0.0006356857484206557
step: 480, loss: 3.001853110617958e-05
step: 490, loss: 3.911008025170304e-05
step: 500, loss: 0.000408057909226045
step: 510, loss: 0.00013673643115907907
step: 520, loss: 0.001900378498248756
step: 530, loss: 2.919561302405782e-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=0.9430523917995445, f1=0.9372460496613996,
best_f1=0.9362292051756008
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 [00:21, 264.38it/s]
load_f1 = 0.9437993497445425

```

```
real_f1 = 0.9447282861124013
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:16, 258.80it/s]
```