

GLow_30_1_2_roberta

August 11, 2022

0.1 Install packages

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 17.00 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.2 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 58.1 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 30.4 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 5.3 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 67.9 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 53.7 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 5.30 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 27.6 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
|                                     | 195 kB 76.6 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 6.9 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.6 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 64.1 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 6.1 MB/s
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Requirement already satisfied: 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: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Collecting blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
    | 3.7 MB 53.4 MB/s
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 77.2 MB/s
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 46.3 MB/s
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.7/dist-packages (from sklearn==0.0->-r requirements.txt
(line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-

```

```

packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    |                                     | 880 kB 11.0 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 46.1 MB/s
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
    Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449922
    sha256=f69a5bc1622d238090b750776764c4622a52ad7eca4f4c9382d487e7f0c310cf
    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=fd11bf7885d773c462f23010b58770e0873f7445dc8db80dd6dfe151cc87d4e6
  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 | 7.24 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-0ey7bfqm
Created temporary directory: /tmp/pip-req-tracker-nb32jcqg
Initialized build tracking at /tmp/pip-req-tracker-nb32jcqg
Created build tracker: /tmp/pip-req-tracker-nb32jcqg
Entered build tracker: /tmp/pip-req-tracker-nb32jcqg
```

```

Created temporary directory: /tmp/pip-install-l75nckfo
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-lkvgebgh
  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-
nb32jcqg'
  Running setup.py (path:/tmp/pip-req-build-lkvgebgh/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-cltc08fp
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-cltc08fp/apex.egg-info
  writing /tmp/pip-pip-egg-info-cltc08fp/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-cltc08fp/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-cltc08fp/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-cltc08fp/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-cltc08fp/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-lkvgebgh 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-nb32jcqg'
Created temporary directory: /tmp/pip-unpack-2xf7tz91
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-8bkkz96g
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-8bkkz96g
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = """/tmp/pip-req-build-lkvgebgh/setup.py""';

```



```
__file__='''/tmp/pip-req-build-lkvgebgh/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-8bkkz96g
```

```
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/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.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/amp.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
creating build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
creating build/lib/apex/normalization
```

```

copying apex/normalization/__init__.py -> build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
creating build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
creating build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
creating build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
creating build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
creating build/lib/apex/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/amp/lists
copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
creating build/lib/apex/contrib/xentropy
copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
creating build/lib/apex/contrib/multihead_attn
copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->

```

```

build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_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/self_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/mask_softmax_dropout_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/fast_encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  creating 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/__init__.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
  creating build/lib/apex/contrib/clip_grad
  copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
  copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
  creating build/lib/apex/contrib/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/transducer
  copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
  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/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
    copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/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/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    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/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
    creating build/lib/apex/transformer/amp
    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/layers
    copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
    copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
    creating build/lib/apex/transformer/functional
    copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
    copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_parallel
    creating build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.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/layers.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.py ->

```

```

build/lib/apex/transformer/tensor_parallel
  creating build/lib/apex/transformer/testing
  copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/global_vars.py ->
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/arguments.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
  creating build/lib/apex/transformer/_data
  copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
  copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
  creating build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py -> 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_pipeline_parallel_without_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipeline_parallel.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
  installing to build/bdist.linux-x86_64/wheel
  running install
  running install_lib
  creating build/bdist.linux-x86_64
  creating build/bdist.linux-x86_64/wheel
  creating build/bdist.linux-x86_64/wheel/apex
  creating build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/scaler.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/handle.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/_amp_state.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/wrap.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/_process_optimizer.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/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/utils.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
  copying build/lib/apex/amp/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  creating 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_adagrad.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_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.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
  creating build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py

```

```

-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/__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/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn.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
    creating 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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_sgd.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/clip_grad
    copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib

```



```

    creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/transducer.py ->
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
    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/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
    copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating build/bdist.linux-x86_64/wheel/apex/contrib/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/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck.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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/__init__.py ->
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
    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/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/multiproc.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/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/grad_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    creating 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
    copying build/lib/apex/transformer/layers/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating

```

```

build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_without_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.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/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/microbatches.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/cross_entropy.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/memory.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/layers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  creating 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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing

```

```

    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/utils.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/_data
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/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
    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-8bkkz96g/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=cf9befd191a460c5176de19968487168d899c7524c66eb94c5457b718ab4d7dd
Stored in directory: /tmp/pip-ephem-wheel-cache-0ey7bfqm/wheels/6b/22/a2/16655
26ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-nb32jcqg'
/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 32.1 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                                     | 3.9 MB 39.8 MB/s
Collecting botocore==1.27.48
  Downloading botocore-1.27.48-py3-none-any.whl (9.0 MB)
    |                                     | 9.0 MB 59.4 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                                     | 548 kB 71.9 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                                     | 79 kB 8.5 MB/s
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.48->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==1.27.48->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-1.25.48 botocore-1.27.48 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 18 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: 1014, done.
remote: Counting objects: 100% (84/84), done.
remote: Compressing objects: 100% (74/74), done.
remote: Total 1014 (delta 27), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (1014/1014), 254.12 MiB | 15.59 MiB/s, done.
Resolving deltas: 100% (611/611), done.
Checking out files: 100% (1284/1284), 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/GLow_30_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 roberta \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da attr_del \
  --summarize

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

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

Downloading: 100% 481/481 [00:00<00:00, 461kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 810kB/s]

Downloading: 100% 456k/456k [00:00<00:00, 514kB/s]

Downloading: 100% 501M/501M [00:10<00:00, 49.9MB/s]

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

Defaults for this optimization level are:

enabled : True

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled            : True
opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4331464469432831
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.32000000000000006, f1=0.2553191489361702,
best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5009341835975647
/usr/local/lib/python3.7/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=0.2828282828282828, f1=0.2692307692307693,
best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44867485761642456
/usr/local/lib/python3.7/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=0.3010752688172043, f1=0.23913043478260868,
best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31462976336479187
/usr/local/lib/python3.7/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=0.288659793814433, f1=0.27184466019417475,
best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3513762354850769
/usr/local/lib/python3.7/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=0.28571428571428575, f1=0.26, best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2661820948123932
/usr/local/lib/python3.7/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=0.2826086956521739, f1=0.2745098039215686,
best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.663905680179596
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.27906976744186046, f1=0.26262626262626265,
best_f1=0.2553191489361702

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5434378981590271
/usr/local/lib/python3.7/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=0.288659793814433, f1=0.27184466019417475,
best_f1=0.2553191489361702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42372992634773254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3783783783783784, f1=0.31578947368421056,
best_f1=0.31578947368421056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44713807106018066
/usr/local/lib/python3.7/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=0.30303030303030304, f1=0.2692307692307692,
best_f1=0.31578947368421056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4340536594390869
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.38596491228070184, f1=0.3283582089552239,
best_f1=0.3283582089552239
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3671049475669861

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.40740740740740744, f1=0.3582089552238806,
best_f1=0.3582089552238806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35708874464035034
/usr/local/lib/python3.7/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=0.38095238095238093, f1=0.37037037037037035,
best_f1=0.3582089552238806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32293468713760376
/usr/local/lib/python3.7/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=0.38095238095238093, f1=0.37037037037037035,
best_f1=0.3582089552238806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3732582628726959
/usr/local/lib/python3.7/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=0.38095238095238093, f1=0.37037037037037035,
best_f1=0.3582089552238806
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, 132990.13it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.3333333333333333
real_f1 = 0.391304347826087
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 145.19it/s]

```

1.1.2 DBLP-ACM - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5636836290359497
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.43069854378700256
step: 20, loss: 0.513922929763794
step: 30, loss: 0.3080146908760071
step: 40, loss: 0.3103317618370056
step: 50, loss: 0.5607637763023376
step: 60, loss: 0.49626070261001587
step: 70, loss: 0.374178946018219
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.1944076418876648
step: 90, loss: 0.14213982224464417
step: 100, loss: 0.2146715223789215
step: 110, loss: 0.023635894060134888
step: 120, loss: 0.07705608755350113
step: 130, loss: 0.1254740059375763
step: 140, loss: 0.016441432759165764
step: 150, loss: 0.17078185081481934
step: 160, loss: 0.01830880157649517
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.023319581523537636
step: 180, loss: 0.05590269714593887
step: 190, loss: 0.021930508315563202
step: 200, loss: 0.022740643471479416
step: 210, loss: 0.1423983871936798
step: 220, loss: 0.41746336221694946
step: 230, loss: 0.00829571858048439
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9797297297297298, f1=0.9806598407281, best_f1=0.9806598407281
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0934344083070755
step: 10, loss: 0.1149289458990097
step: 20, loss: 0.015596337616443634
step: 30, loss: 0.005982834845781326
step: 40, loss: 0.06713251024484634
step: 50, loss: 0.005653634667396545
step: 60, loss: 0.006081579253077507
step: 70, loss: 0.005692570935934782

```

```

step: 80, loss: 0.01865246705710888
step: 90, loss: 0.02881813794374466
step: 100, loss: 0.003101848531514406
step: 110, loss: 0.003982567694038153
step: 120, loss: 0.009093660861253738
step: 130, loss: 0.002152681117877364
step: 140, loss: 0.002555637853220105
step: 150, loss: 0.09665972739458084
step: 160, loss: 0.0020001919474452734
step: 170, loss: 0.006850256584584713
step: 180, loss: 0.0040092081762850285
step: 190, loss: 0.009131614118814468
step: 200, loss: 0.01503078918904066
step: 210, loss: 0.0017183322925120592
step: 220, loss: 0.0014021436218172312
step: 230, loss: 0.0011410336010158062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921436588103255, f1=0.9853107344632768,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01952669583261013
step: 10, loss: 0.0012949268566444516
step: 20, loss: 0.011695606634020805
step: 30, loss: 0.0003469569201115519
step: 40, loss: 0.005650754552334547
step: 50, loss: 0.025207892060279846
step: 60, loss: 0.0008152239606715739
step: 70, loss: 0.00048291831626556814
step: 80, loss: 0.008541770279407501
step: 90, loss: 0.0003689599398057908
step: 100, loss: 0.004371258430182934
step: 110, loss: 0.045390233397483826
step: 120, loss: 0.00041488101123832166
step: 130, loss: 0.008538979105651379
step: 140, loss: 0.0003647358389571309
step: 150, loss: 0.0007806390058249235
step: 160, loss: 0.0019964936655014753
step: 170, loss: 0.002467723796144128
step: 180, loss: 0.014069276861846447
step: 190, loss: 0.09527404606342316

```

```

step: 200, loss: 0.011318385601043701
step: 210, loss: 0.00171396613586694
step: 220, loss: 0.025595778599381447
step: 230, loss: 0.007389153819531202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9865168539325843, f1=0.9819819819819819,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002440342213958502
step: 10, loss: 0.008252641186118126
step: 20, loss: 0.0018299923976883292
step: 30, loss: 0.000778384564910084
step: 40, loss: 0.020243573933839798
step: 50, loss: 0.00799773819744587
step: 60, loss: 0.0007297112024389207
step: 70, loss: 0.00043182290391996503
step: 80, loss: 0.00036159378942102194
step: 90, loss: 0.0009288639412261546
step: 100, loss: 0.057467710226774216
step: 110, loss: 0.0004024609224870801
step: 120, loss: 0.008389398455619812
step: 130, loss: 0.0661599189043045
step: 140, loss: 0.0012534355046227574
step: 150, loss: 0.0011560515267774463
step: 160, loss: 0.008173671551048756
step: 170, loss: 0.0013882371131330729
step: 180, loss: 0.04516893997788429
step: 190, loss: 0.0020500347018241882
step: 200, loss: 0.013923181220889091
step: 210, loss: 0.0005518165417015553
step: 220, loss: 0.000408862077165395
step: 230, loss: 0.03710102289915085
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.980963045912654,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009136219508945942
step: 10, loss: 0.00120366585906595
step: 20, loss: 0.000935705378651619
step: 30, loss: 0.0006995750009082258

```

```

step: 40, loss: 0.0019429567037150264
step: 50, loss: 0.001483504893258214
step: 60, loss: 0.01905268058180809
step: 70, loss: 0.0014636935666203499
step: 80, loss: 0.0997706800699234
step: 90, loss: 0.22843872010707855
step: 100, loss: 0.00037724681897088885
step: 110, loss: 0.0018034311942756176
step: 120, loss: 0.007834569551050663
step: 130, loss: 0.0009417987894266844
step: 140, loss: 0.00046107705566100776
step: 150, loss: 0.007089603692293167
step: 160, loss: 0.00026611192151904106
step: 170, loss: 0.06134617701172829
step: 180, loss: 0.0012986185029149055
step: 190, loss: 0.008961399085819721
step: 200, loss: 0.0024369312450289726
step: 210, loss: 0.0019396889256313443
step: 220, loss: 0.009050303138792515
step: 230, loss: 0.002208676654845476
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9820224719101124,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005513879004865885
step: 10, loss: 0.0015689695719629526
step: 20, loss: 0.0013925032690167427
step: 30, loss: 0.0007621293771080673
step: 40, loss: 0.00042577661224640906
step: 50, loss: 0.002328814473003149
step: 60, loss: 0.004564226139336824
step: 70, loss: 0.00048433500342071056
step: 80, loss: 0.002775057451799512
step: 90, loss: 0.01907210610806942
step: 100, loss: 0.0007869155379012227
step: 110, loss: 0.0025647366419434547
step: 120, loss: 0.00033019654802046716
step: 130, loss: 0.00033897580578923225
step: 140, loss: 0.0029003506060689688
step: 150, loss: 0.000366113061318174
step: 160, loss: 0.008599048480391502
step: 170, loss: 0.0003805147425737232
step: 180, loss: 0.00014863445539958775
step: 190, loss: 0.00025960011407732964

```

```

step: 200, loss: 0.0013656769879162312
step: 210, loss: 0.000670926587190479
step: 220, loss: 0.0038198246620595455
step: 230, loss: 0.005951540544629097
/usr/local/lib/python3.7/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=0.9778270509977827,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017134519293904305
step: 10, loss: 0.0007860591867938638
step: 20, loss: 0.0003669930447358638
step: 30, loss: 0.0002579804859124124
step: 40, loss: 0.0010132978204637766
step: 50, loss: 0.0007936623296700418
step: 60, loss: 0.0005967579199932516
step: 70, loss: 0.0005251431139186025
step: 80, loss: 0.0007408235687762499
step: 90, loss: 0.0003714611229952425
step: 100, loss: 0.00044758402509614825
step: 110, loss: 0.00038148349267430604
step: 120, loss: 0.0003266257990617305
step: 130, loss: 0.001133798505179584
step: 140, loss: 0.00036884009023196995
step: 150, loss: 0.0065875668078660965
step: 160, loss: 0.0001435075100744143
step: 170, loss: 0.005015556700527668
step: 180, loss: 0.0017232041573151946
step: 190, loss: 0.0006413991213776171
step: 200, loss: 0.0017957298550754786
step: 210, loss: 0.004229755140841007
step: 220, loss: 0.00045845392742194235
step: 230, loss: 0.0006453878013417125
/usr/local/lib/python3.7/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=0.9832026875699889,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005729975528083742
step: 10, loss: 0.0015380831900984049
step: 20, loss: 0.0016404666239395738
step: 30, loss: 0.0017168757040053606

```



```

step: 40, loss: 0.0077666365541517735
step: 50, loss: 0.01044660247862339
step: 60, loss: 0.00083203858230263
step: 70, loss: 0.0003288064617663622
step: 80, loss: 0.005285758525133133
step: 90, loss: 0.0008087293244898319
step: 100, loss: 0.0007077318732626736
step: 110, loss: 0.018262667581439018
step: 120, loss: 0.00030439384863711894
step: 130, loss: 0.0007672672509215772
step: 140, loss: 0.0018436408136039972
step: 150, loss: 0.049270521849393845
step: 160, loss: 0.0004067840927746147
step: 170, loss: 0.00695757195353508
step: 180, loss: 0.0003648429410532117
step: 190, loss: 0.0005547809996642172
step: 200, loss: 0.06246250495314598
step: 210, loss: 0.0019145311089232564
step: 220, loss: 0.0010649396572262049
step: 230, loss: 0.0006880909204483032
/usr/local/lib/python3.7/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=0.9822222222222222,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042967277113348246
step: 10, loss: 0.0006495500565506518
step: 20, loss: 0.0007380139431916177
step: 30, loss: 0.000704994541592896
step: 40, loss: 0.0005485265282914042
step: 50, loss: 0.0006836080574430525
step: 60, loss: 0.0005126541364006698
step: 70, loss: 0.11372467875480652
step: 80, loss: 0.000297837977996096
step: 90, loss: 0.06908443570137024
step: 100, loss: 0.0005675890133716166
step: 110, loss: 0.0004952551680617034
step: 120, loss: 0.023844702169299126
step: 130, loss: 0.0006804590811952949
step: 140, loss: 0.0007022889330983162
step: 150, loss: 0.0033040123526006937
step: 160, loss: 0.0008476969669573009
step: 170, loss: 0.00041072044405154884
step: 180, loss: 0.0006977714947424829
step: 190, loss: 0.0002605762565508485

```

```

step: 200, loss: 0.00039015518268570304
step: 210, loss: 0.0008465402061119676
step: 220, loss: 0.0005048968596383929
step: 230, loss: 0.00028692462365143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921612541993281, f1=0.9821428571428571,
best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004699409764725715
step: 10, loss: 0.0007939409697428346
step: 20, loss: 0.0004937727935612202
step: 30, loss: 0.00029421172803267837
step: 40, loss: 0.001576487091369927
step: 50, loss: 0.00034454435808584094
step: 60, loss: 0.0013583017280325294
step: 70, loss: 0.002316203899681568
step: 80, loss: 0.0009394121589139104
step: 90, loss: 0.0005180130829103291
step: 100, loss: 0.00042636407306417823
step: 110, loss: 0.0006295401835814118
step: 120, loss: 0.00028609350556507707
step: 130, loss: 0.0003824271261692047
step: 140, loss: 0.00030697902548126876
step: 150, loss: 0.00042745377868413925
step: 160, loss: 0.00016843312187120318
step: 170, loss: 0.0003289133310317993
step: 180, loss: 0.0005185601767152548
step: 190, loss: 0.00041302863974124193
step: 200, loss: 0.0005825061234645545
step: 210, loss: 0.0004820358590222895
step: 220, loss: 0.00027193102869205177
step: 230, loss: 0.00020519556710496545
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.9843749999999999,
best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00018985228962264955
step: 10, loss: 0.000431023130659014
step: 20, loss: 0.0003494042612146586
step: 30, loss: 0.0002906131849158555
step: 40, loss: 6.034219768480398e-05
step: 50, loss: 0.0001344540505670011
step: 60, loss: 0.011202649213373661
step: 70, loss: 0.0019135293550789356
step: 80, loss: 0.0007382422918453813
step: 90, loss: 0.20896124839782715
step: 100, loss: 0.0005113906809128821
step: 110, loss: 0.00036392174661159515
step: 120, loss: 0.0006448577041737735
step: 130, loss: 0.0005625479971058667
step: 140, loss: 0.012525994330644608
step: 150, loss: 0.0007678517140448093
step: 160, loss: 0.0010177791118621826
step: 170, loss: 0.0013455325970426202
step: 180, loss: 0.0011019816156476736
step: 190, loss: 0.0008643295732326806
step: 200, loss: 0.0010010561672970653
step: 210, loss: 0.0008754449663683772
step: 220, loss: 0.0009559168829582632
step: 230, loss: 0.000494087697006762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9910313901345291, f1=0.9854423292273236,
best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009315382922068238
step: 10, loss: 0.0005642649484798312
step: 20, loss: 0.04532548412680626
step: 30, loss: 0.045050498098134995
step: 40, loss: 0.0009916916023939848
step: 50, loss: 0.0007910010754130781
step: 60, loss: 0.0008568699704483151
step: 70, loss: 0.0006183389341458678
step: 80, loss: 0.00028265287983231246
step: 90, loss: 0.0026842150837183
step: 100, loss: 0.00041957030771300197
step: 110, loss: 0.00026856025215238333
step: 120, loss: 0.00045762176159769297
step: 130, loss: 0.00048285420052707195
step: 140, loss: 0.0005821561790071428
step: 150, loss: 0.0005865016719326377

```

```

step: 160, loss: 0.00048681884072721004
step: 170, loss: 0.000499196641612798
step: 180, loss: 0.00036791712045669556
step: 190, loss: 0.0016836603172123432
step: 200, loss: 0.0003412653459236026
step: 210, loss: 0.000504685565829277
step: 220, loss: 0.01279835682362318
step: 230, loss: 0.0023310952819883823
/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9843749999999999,
best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003975949657615274
step: 10, loss: 0.0006047877832315862
step: 20, loss: 0.0009626976097933948
step: 30, loss: 0.00046217910130508244
step: 40, loss: 0.000977497547864914
step: 50, loss: 0.0011463086120784283
step: 60, loss: 0.0006374206277541816
step: 70, loss: 0.0005027705919928849
step: 80, loss: 0.0005315483431331813
step: 90, loss: 0.0006715963827446103
step: 100, loss: 0.001206598011776805
step: 110, loss: 0.0012335567735135555
step: 120, loss: 0.0009680954972282052
step: 130, loss: 0.0010188821470364928
step: 140, loss: 0.0005813906318508089
step: 150, loss: 0.00017532712081447244
step: 160, loss: 0.0025833502877503633
step: 170, loss: 0.0009308304288424551
step: 180, loss: 0.04538977891206741
step: 190, loss: 0.0007764690672047436
step: 200, loss: 0.00018426212773192674
step: 210, loss: 0.0006769606843590736
step: 220, loss: 0.0006782260606996715
step: 230, loss: 0.00046652305172756314
/usr/local/lib/python3.7/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=0.9921259842519685, f1=0.9854423292273236,
best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00040059664752334356
step: 10, loss: 0.0003390883794054389
step: 20, loss: 0.00040797083056531847
step: 30, loss: 0.0007457262836396694
step: 40, loss: 0.0004989856388419867
step: 50, loss: 0.00041134635102935135
step: 60, loss: 0.0005707741365768015
step: 70, loss: 0.0003867428167723119
step: 80, loss: 0.0004729698412120342
step: 90, loss: 0.0009163948125205934
step: 100, loss: 0.0007133355247788131
step: 110, loss: 0.0010916151804849505
step: 120, loss: 0.00022067571990191936
step: 130, loss: 0.001007341081276536
step: 140, loss: 0.0004914972232654691
step: 150, loss: 0.0002613381075207144
step: 160, loss: 0.0005681032780557871
step: 170, loss: 0.0006125162472017109
step: 180, loss: 0.0007285886676982045
step: 190, loss: 0.0004828832461498678
step: 200, loss: 0.000733414082787931
step: 210, loss: 0.000456993468105793
step: 220, loss: 0.0005984374438412488
step: 230, loss: 0.0002203651238232851
/usr/local/lib/python3.7/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=0.9921259842519685, f1=0.9876543209876544,
best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025633830577135086
step: 10, loss: 0.0005683837225660682
step: 20, loss: 0.0005546712200157344
step: 30, loss: 0.0006478952127508819
step: 40, loss: 0.00042307787225581706
step: 50, loss: 0.0005240490427240729
step: 60, loss: 0.02721014991402626
step: 70, loss: 0.000460839772131294
step: 80, loss: 0.00044102611718699336
step: 90, loss: 0.0004376518481876701
step: 100, loss: 0.000312309741275385
step: 110, loss: 0.0005916512454859912
step: 120, loss: 0.02744237519800663
step: 130, loss: 0.0003454524266999215
step: 140, loss: 0.0008016053470782936
step: 150, loss: 0.0004958981298841536

```

```

step: 160, loss: 0.017098715528845787
step: 170, loss: 0.00022963975789025426
step: 180, loss: 0.0004943585372529924
step: 190, loss: 0.000992717337794602
step: 200, loss: 0.0004975451738573611
step: 210, loss: 0.029176514595746994
step: 220, loss: 0.0005503084976226091
step: 230, loss: 0.000587887829169631
/usr/local/lib/python3.7/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=0.9921259842519685, f1=0.9843400447427293,
best_f1=0.9821428571428571
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 [00:16, 148.59it/s]
load_f1 = 0.9887640449438202
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 [00:32, 134.14it/s]
```

1.1.3 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : O2
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.669550359249115
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4195062816143036
step: 20, loss: 0.2939992845058441
step: 30, loss: 0.3399136960506439
step: 40, loss: 0.4081266522407532
step: 50, loss: 0.5715963840484619
step: 60, loss: 0.34054669737815857
step: 70, loss: 0.40610283613204956
step: 80, loss: 0.20530211925506592
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.20161287486553192
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 100, loss: 0.3761182427406311
step: 110, loss: 0.12816035747528076
step: 120, loss: 0.10695955902338028
step: 130, loss: 0.11616037040948868
step: 140, loss: 0.20506338775157928
step: 150, loss: 0.11797469854354858
step: 160, loss: 0.18301571905612946
step: 170, loss: 0.060318805277347565
step: 180, loss: 0.039870038628578186
step: 190, loss: 0.05985328555107117
step: 200, loss: 0.017971906810998917
step: 210, loss: 0.10648854821920395
step: 220, loss: 0.20889177918434143
step: 230, loss: 0.15768173336982727
step: 240, loss: 0.01876257173717022
step: 250, loss: 0.0823723077740479
step: 260, loss: 0.16984528303146362
step: 270, loss: 0.19175264239311218

```



```

step: 280, loss: 0.06598357856273651
step: 290, loss: 0.09944842755794525
step: 300, loss: 0.017650321125984192
step: 310, loss: 0.1269105076789856
step: 320, loss: 0.05512382462620735
step: 330, loss: 0.08852215111255646
step: 340, loss: 0.20543991029262543
step: 350, loss: 0.1516328603029251
step: 360, loss: 0.05688728764653206
step: 370, loss: 0.1314474493265152
step: 380, loss: 0.14652742445468903
step: 390, loss: 0.056952305138111115
step: 400, loss: 0.043547432869672775
step: 410, loss: 0.30195939540863037
step: 420, loss: 0.06389030814170837
step: 430, loss: 0.01623852550983429
step: 440, loss: 0.04714839905500412
step: 450, loss: 0.068548783659935
step: 460, loss: 0.008600003086030483
step: 470, loss: 0.027861641719937325
step: 480, loss: 0.19250690937042236
step: 490, loss: 0.18659934401512146
step: 500, loss: 0.020418405532836914
step: 510, loss: 0.020564578473567963
step: 520, loss: 0.050684235990047455
step: 530, loss: 0.012641050852835178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9332096474953617, f1=0.9355586462679647,
best_f1=0.9355586462679647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0687556192278862
step: 10, loss: 0.12754717469215393
step: 20, loss: 0.05493040010333061
step: 30, loss: 0.09992367029190063
step: 40, loss: 0.06392578035593033
step: 50, loss: 0.044668033719062805
step: 60, loss: 0.017962567508220673
step: 70, loss: 0.01628785952925682
step: 80, loss: 0.037765856832265854
step: 90, loss: 0.008441504091024399

```

```
step: 100, loss: 0.10467593371868134
step: 110, loss: 0.02469269186258316
step: 120, loss: 0.2304845005273819
step: 130, loss: 0.012612725608050823
step: 140, loss: 0.0401589460670948
step: 150, loss: 0.010674362070858479
step: 160, loss: 0.026987671852111816
step: 170, loss: 0.04285478591918945
step: 180, loss: 0.012584289535880089
step: 190, loss: 0.02114001102745533
step: 200, loss: 0.08666857331991196
step: 210, loss: 0.03909960016608238
step: 220, loss: 0.0018173570279031992
step: 230, loss: 0.024095207452774048
step: 240, loss: 0.014724023640155792
step: 250, loss: 0.019231101498007774
step: 260, loss: 0.008825830183923244
step: 270, loss: 0.040366917848587036
step: 280, loss: 0.06139587610960007
step: 290, loss: 0.010247462429106236
step: 300, loss: 0.022158939391374588
step: 310, loss: 0.12712661921977997
step: 320, loss: 0.07554920017719269
step: 330, loss: 0.07542437314987183
step: 340, loss: 0.01337869931012392
step: 350, loss: 0.003734284546226263
step: 360, loss: 0.08901064097881317
step: 370, loss: 0.008839966729283333
step: 380, loss: 0.19304247200489044
step: 390, loss: 0.008623073808848858
step: 400, loss: 0.02693350799381733
step: 410, loss: 0.01432715356349945
step: 420, loss: 0.1936599612236023
step: 430, loss: 0.23046515882015228
step: 440, loss: 0.039347678422927856
step: 450, loss: 0.03466096520423889
step: 460, loss: 0.05884310603141785
step: 470, loss: 0.08827514946460724
step: 480, loss: 0.01726638339459896
step: 490, loss: 0.055991046130657196
step: 500, loss: 0.01454133354127407
step: 510, loss: 0.017566172406077385
step: 520, loss: 0.4646325707435608
step: 530, loss: 0.035747211426496506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9511970534069981, f1=0.9475620975160993,

best_f1=0.9475620975160993

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

An input tensor was not cuda.

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

step: 0, loss: 0.16356390714645386

step: 10, loss: 0.042384881526231766

step: 20, loss: 0.0052140457555651665

step: 30, loss: 0.05095289647579193

step: 40, loss: 0.06453626602888107

step: 50, loss: 0.0034347069449722767

step: 60, loss: 0.022769033908843994

step: 70, loss: 0.018438735976815224

step: 80, loss: 0.002785841468721628

step: 90, loss: 0.00194695929531008

step: 100, loss: 0.04196985438466072

step: 110, loss: 0.040773507207632065

step: 120, loss: 0.1142403781414032

step: 130, loss: 0.08453156054019928

step: 140, loss: 0.011224950663745403

step: 150, loss: 0.012097076512873173

step: 160, loss: 0.04621583968400955

step: 170, loss: 0.0011180833680555224

step: 180, loss: 0.032590966671705246

step: 190, loss: 0.006960522383451462

step: 200, loss: 0.06591156870126724

step: 210, loss: 0.0651470273733139

step: 220, loss: 0.14765852689743042

step: 230, loss: 0.021389957517385483

step: 240, loss: 0.03055400215089321

step: 250, loss: 0.01744457706809044

step: 260, loss: 0.05670371651649475

step: 270, loss: 0.001551260007545352

step: 280, loss: 0.006185026839375496

step: 290, loss: 0.003927480895072222

step: 300, loss: 0.08949033915996552

step: 310, loss: 0.026963869109749794

step: 320, loss: 0.11065874993801117

step: 330, loss: 0.014538195915520191

step: 340, loss: 0.009643030352890491

step: 350, loss: 0.17270280420780182

step: 360, loss: 0.026336712762713432

step: 370, loss: 0.018356705084443092

step: 380, loss: 0.05364491418004036

step: 390, loss: 0.007601856254041195

```

step: 400, loss: 0.07735573500394821
step: 410, loss: 0.08098220080137253
step: 420, loss: 0.00943084992468357
step: 430, loss: 0.009996398352086544
step: 440, loss: 0.08394671976566315
step: 450, loss: 0.09143244475126266
step: 460, loss: 0.14800699055194855
step: 470, loss: 0.04608476907014847
step: 480, loss: 0.02943829819560051
step: 490, loss: 0.04709787294268608
step: 500, loss: 0.007004184648394585
step: 510, loss: 0.043796613812446594
step: 520, loss: 0.0013310223584994674
step: 530, loss: 0.014480902813374996
/usr/local/lib/python3.7/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=0.9480642115203022, f1=0.9459715639810427,
best_f1=0.9475620975160993
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02340666577219963
step: 10, loss: 0.0019690911285579205
step: 20, loss: 0.010420096106827259
step: 30, loss: 0.13131825625896454
step: 40, loss: 0.03305281326174736
step: 50, loss: 0.04708680510520935
step: 60, loss: 0.016615867614746094
step: 70, loss: 0.025254931300878525
step: 80, loss: 0.004709798842668533
step: 90, loss: 0.15441565215587616
step: 100, loss: 0.004396635107696056
step: 110, loss: 0.018638715147972107
step: 120, loss: 0.0005229129455983639
step: 130, loss: 0.11067253351211548
step: 140, loss: 0.01591203361749649
step: 150, loss: 0.015329516492784023
step: 160, loss: 0.006059572566300631
step: 170, loss: 0.0648660808801651
step: 180, loss: 0.10279133915901184
step: 190, loss: 0.053017303347587585
step: 200, loss: 0.020510682836174965
step: 210, loss: 0.004000226967036724
step: 220, loss: 0.001284438301809132
step: 230, loss: 0.009835384786128998
step: 240, loss: 0.03910485655069351
step: 250, loss: 0.034824103116989136

```

```

step: 260, loss: 0.017385758459568024
step: 270, loss: 0.01623055338859558
step: 280, loss: 0.0030514320824295282
step: 290, loss: 0.1303287148475647
step: 300, loss: 0.007417095825076103
step: 310, loss: 0.011213990859687328
step: 320, loss: 0.07033386081457138
step: 330, loss: 0.035998694598674774
step: 340, loss: 0.006538609508424997
step: 350, loss: 0.05772673338651657
step: 360, loss: 0.012533855624496937
step: 370, loss: 0.004122977145016193
step: 380, loss: 0.017292698845267296
step: 390, loss: 0.00024471842334605753
step: 400, loss: 0.003950178157538176
step: 410, loss: 0.0024919966235756874
step: 420, loss: 0.0034746762830764055
step: 430, loss: 0.004015001934021711
step: 440, loss: 0.006662419997155666
step: 450, loss: 0.026603126898407936
step: 460, loss: 0.01705736294388771
step: 470, loss: 0.001014632172882557
step: 480, loss: 0.01783330738544464
step: 490, loss: 0.0024068085476756096
step: 500, loss: 0.024160325527191162
step: 510, loss: 0.04135288670659065
step: 520, loss: 0.006965420208871365
step: 530, loss: 0.08556749671697617
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9533487297921477, f1=0.9486940298507462,
best_f1=0.9486940298507462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04467804357409477
step: 10, loss: 0.057563770562410355
step: 20, loss: 0.04172641783952713
step: 30, loss: 0.033350422978401184
step: 40, loss: 0.0003324617864564061
step: 50, loss: 0.17728659510612488
step: 60, loss: 0.12306834012269974
step: 70, loss: 0.0005513170617632568

```

```
step: 80, loss: 0.022589154541492462
step: 90, loss: 0.06275233626365662
step: 100, loss: 0.058683570474386215
step: 110, loss: 0.0027004978619515896
step: 120, loss: 0.1933160275220871
step: 130, loss: 0.011958044022321701
step: 140, loss: 0.00862779002636671
step: 150, loss: 0.020381532609462738
step: 160, loss: 0.0009793051285669208
step: 170, loss: 0.12498854100704193
step: 180, loss: 0.04931693896651268
step: 190, loss: 0.003401403082534671
step: 200, loss: 0.004467909224331379
step: 210, loss: 0.0014240809250622988
step: 220, loss: 0.00148982263635844
step: 230, loss: 0.000863684166688472
step: 240, loss: 0.0019918703474104404
step: 250, loss: 0.10464134812355042
step: 260, loss: 0.0005124580347910523
step: 270, loss: 0.20381274819374084
step: 280, loss: 0.013151859864592552
step: 290, loss: 0.025526059791445732
step: 300, loss: 0.030765188857913017
step: 310, loss: 0.023857057094573975
step: 320, loss: 0.053111184388399124
step: 330, loss: 0.0006541286129504442
step: 340, loss: 0.035510797053575516
step: 350, loss: 0.0008110931958071887
step: 360, loss: 0.0002520614070817828
step: 370, loss: 0.0011655777925625443
step: 380, loss: 0.0008979398990049958
step: 390, loss: 0.001857293420471251
step: 400, loss: 0.002274904865771532
step: 410, loss: 0.09348899126052856
step: 420, loss: 0.14731566607952118
step: 430, loss: 0.04085769131779671
step: 440, loss: 0.0008031772449612617
step: 450, loss: 0.007901903241872787
step: 460, loss: 0.01398627832531929
step: 470, loss: 0.04543383792042732
step: 480, loss: 0.015551600605249405
step: 490, loss: 0.020160026848316193
step: 500, loss: 0.053594935685396194
step: 510, loss: 0.0025136431213468313
step: 520, loss: 0.09693863987922668
step: 530, loss: 0.021171126514673233
/usr/local/lib/python3.7/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=0.9524680073126143, f1=0.9537757437070938,
best_f1=0.9486940298507462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015559767372906208
step: 10, loss: 0.0001460238272557035
step: 20, loss: 0.00015620168414898217
step: 30, loss: 0.0002827758726198226
step: 40, loss: 0.011226539500057697
step: 50, loss: 0.039318423718214035
step: 60, loss: 0.0009237237391062081
step: 70, loss: 0.016153786331415176
step: 80, loss: 0.0016893476713448763
step: 90, loss: 0.018057027831673622
step: 100, loss: 0.0025924036744982004
step: 110, loss: 0.006013024598360062
step: 120, loss: 0.028325367718935013
step: 130, loss: 0.0008111554197967052
step: 140, loss: 0.0070214299485087395
step: 150, loss: 0.0022917368914932013
step: 160, loss: 0.029494550079107285
step: 170, loss: 0.0034888293594121933
step: 180, loss: 0.0027600079774856567
step: 190, loss: 0.1808980256319046
step: 200, loss: 0.008971295319497585
step: 210, loss: 0.0070111677050590515
step: 220, loss: 0.009839109145104885
step: 230, loss: 0.01714484766125679
step: 240, loss: 0.13400374352931976
step: 250, loss: 0.020083114504814148
step: 260, loss: 0.0033937660045921803
step: 270, loss: 0.0037321054842323065
step: 280, loss: 0.004512095358222723
step: 290, loss: 0.0822344645857811
step: 300, loss: 0.008148560300469398
step: 310, loss: 0.013688342645764351
step: 320, loss: 0.01998247392475605
step: 330, loss: 0.006311905104666948
step: 340, loss: 0.0013903798535466194
step: 350, loss: 0.023430004715919495
step: 360, loss: 0.03808128461241722
step: 370, loss: 0.006418080069124699
step: 380, loss: 0.00016715956735424697
step: 390, loss: 0.008601007983088493
step: 400, loss: 0.05850493162870407
step: 410, loss: 0.00023764913203194737
```

```

step: 420, loss: 0.009324634447693825
step: 430, loss: 0.008406850509345531
step: 440, loss: 0.0008680336177349091
step: 450, loss: 0.23796777427196503
step: 460, loss: 0.0034716606605798006
step: 470, loss: 0.02230020798742771
step: 480, loss: 0.026219090446829796
step: 490, loss: 0.011444872245192528
step: 500, loss: 0.002870525000616908
step: 510, loss: 0.19361354410648346
step: 520, loss: 0.0011149101192131639
step: 530, loss: 0.0049735126085579395
/usr/local/lib/python3.7/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=0.9513822688274547, f1=0.942077549066539,
best_f1=0.9486940298507462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024002513382583857
step: 10, loss: 0.001150991884060204
step: 20, loss: 0.008263758383691311
step: 30, loss: 0.0053455280140042305
step: 40, loss: 0.01960626430809498
step: 50, loss: 0.001224066480062902
step: 60, loss: 0.006094735581427813
step: 70, loss: 0.00031340765417553484
step: 80, loss: 0.0018411902710795403
step: 90, loss: 0.00014481630933005363
step: 100, loss: 0.006927599664777517
step: 110, loss: 0.0002047439629677683
step: 120, loss: 0.0017184971366077662
step: 130, loss: 0.0008099208935163915
step: 140, loss: 0.0007172978948801756
step: 150, loss: 0.001175579964183271
step: 160, loss: 0.0011291262926533818
step: 170, loss: 0.014311378821730614
step: 180, loss: 0.059003379195928574
step: 190, loss: 0.05411148816347122
step: 200, loss: 0.0001410875702276826
step: 210, loss: 0.0022910567931830883
step: 220, loss: 0.00036353952600620687
step: 230, loss: 0.0007825841894373298
step: 240, loss: 0.002132267225533724
step: 250, loss: 0.1011691614985466
step: 260, loss: 0.003306439844891429
step: 270, loss: 9.496649727225304e-05

```



```

step: 280, loss: 0.00041765530477277935
step: 290, loss: 0.0005924816941842437
step: 300, loss: 0.00010839279275387526
step: 310, loss: 0.009972186759114265
step: 320, loss: 0.03820731118321419
step: 330, loss: 0.0034091423731297255
step: 340, loss: 0.029787657782435417
step: 350, loss: 0.004273723345249891
step: 360, loss: 0.003521230071783066
step: 370, loss: 0.04983583092689514
step: 380, loss: 0.029675347730517387
step: 390, loss: 0.009629608131945133
step: 400, loss: 0.015803249552845955
step: 410, loss: 0.0008576668333262205
step: 420, loss: 0.028660036623477936
step: 430, loss: 0.005074556451290846
step: 440, loss: 0.00038388444227166474
step: 450, loss: 0.012397664599120617
step: 460, loss: 0.0028307163156569004
step: 470, loss: 0.12755407392978668
step: 480, loss: 0.005657503381371498
step: 490, loss: 0.0005980920977890491
step: 500, loss: 0.0015261759981513023
step: 510, loss: 0.00021696821204386652
step: 520, loss: 0.00014839658979326487
step: 530, loss: 0.0010932020377367735
/usr/local/lib/python3.7/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=0.9522484932777004, f1=0.952513966480447,
best_f1=0.9486940298507462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025370262563228607
step: 10, loss: 0.010684196837246418
step: 20, loss: 0.0012657255865633488
step: 30, loss: 0.0004954517353326082
step: 40, loss: 0.00011114068911410868
step: 50, loss: 0.00036434835055842996
step: 60, loss: 0.0002223569026682526
step: 70, loss: 0.0003195166646037251
step: 80, loss: 0.09389106184244156
step: 90, loss: 0.00020051162573508918
step: 100, loss: 0.0012830322375521064
step: 110, loss: 0.0015458350535482168
step: 120, loss: 0.00020761150517500937
step: 130, loss: 0.020185561850667

```

```

step: 140, loss: 0.00014154569362290204
step: 150, loss: 0.0029046316631138325
step: 160, loss: 0.00034075399162247777
step: 170, loss: 0.0010963325621560216
step: 180, loss: 0.00598517619073391
step: 190, loss: 0.0010206819279119372
step: 200, loss: 0.0033908269833773375
step: 210, loss: 0.045698996633291245
step: 220, loss: 0.0002827791904564947
step: 230, loss: 0.03219112008810043
step: 240, loss: 0.0008574354578740895
step: 250, loss: 0.0015437123365700245
step: 260, loss: 0.004817355889827013
step: 270, loss: 0.015982620418071747
step: 280, loss: 0.005454178433865309
step: 290, loss: 0.004276665393263102
step: 300, loss: 6.406284956028685e-05
step: 310, loss: 0.0025019622407853603
step: 320, loss: 0.00027582378243096173
step: 330, loss: 0.0006805849261581898
step: 340, loss: 0.02013384737074375
step: 350, loss: 0.002773247193545103
step: 360, loss: 0.007388461846858263
step: 370, loss: 0.017260082066059113
step: 380, loss: 0.02734171412885189
step: 390, loss: 0.0005308273830451071
step: 400, loss: 0.0027274922467768192
step: 410, loss: 0.0006296646315604448
step: 420, loss: 0.00019465519289951771
step: 430, loss: 0.0014742766506969929
step: 440, loss: 0.00037659064400941133
step: 450, loss: 0.0001264343736693263
step: 460, loss: 0.0032185502350330353
step: 470, loss: 0.03804170712828636
step: 480, loss: 0.0008532723877578974
step: 490, loss: 0.004311608150601387
step: 500, loss: 0.00011860524682560936
step: 510, loss: 0.002474299632012844
step: 520, loss: 0.0009233971359208226
step: 530, loss: 0.0006304102716967463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9543761638733707, f1=0.9535315985130112,

```

```
best_f1=0.9535315985130112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001147114162449725
step: 10, loss: 0.010518846102058887
step: 20, loss: 0.0006519252783618867
step: 30, loss: 0.07154390215873718
step: 40, loss: 0.0002380638470640406
step: 50, loss: 0.00014507446030620486
step: 60, loss: 0.0007904127705842257
step: 70, loss: 0.027700791135430336
step: 80, loss: 0.013150237500667572
step: 90, loss: 0.04548831656575203
step: 100, loss: 7.69938196754083e-05
step: 110, loss: 0.0032402167562395334
step: 120, loss: 0.0005177208804525435
step: 130, loss: 0.0006445215549319983
step: 140, loss: 4.9645561375655234e-05
step: 150, loss: 0.001089353347197175
step: 160, loss: 0.0005838056677021086
step: 170, loss: 0.030332330614328384
step: 180, loss: 0.007405344396829605
step: 190, loss: 9.408647019881755e-05
step: 200, loss: 0.0011059982934966683
step: 210, loss: 0.0018570753745734692
step: 220, loss: 0.0008791141444817185
step: 230, loss: 0.0007788304355926812
step: 240, loss: 0.0002816363994497806
step: 250, loss: 0.0004770267696585506
step: 260, loss: 0.004749559331685305
step: 270, loss: 0.004084477201104164
step: 280, loss: 0.01785879209637642
step: 290, loss: 0.00026055911439470947
step: 300, loss: 0.0008275185828097165
step: 310, loss: 0.004619607236236334
step: 320, loss: 0.0006284329574555159
step: 330, loss: 0.000575465674046427
step: 340, loss: 0.00480481144040823
step: 350, loss: 0.00986684113740921
step: 360, loss: 0.0008401017403230071
step: 370, loss: 0.0004876490856986493
step: 380, loss: 0.0005797410267405212
step: 390, loss: 0.00023557418899144977
step: 400, loss: 0.0105748837813735
step: 410, loss: 0.0001111429010052234
step: 420, loss: 0.0005027704173699021
step: 430, loss: 0.002125832485035062
```

```

step: 440, loss: 8.203453035093844e-05
step: 450, loss: 0.05492168292403221
step: 460, loss: 4.733592504635453e-05
step: 470, loss: 0.0002987525367643684
step: 480, loss: 3.214979369658977e-05
step: 490, loss: 0.0008258220041170716
step: 500, loss: 0.030203012749552727
step: 510, loss: 0.00040444309706799686
step: 520, loss: 0.004958549980074167
step: 530, loss: 0.01737804152071476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9553201289728237, f1=0.9532967032967032,
best_f1=0.9532967032967032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030774062033742666
step: 10, loss: 0.0011226017959415913
step: 20, loss: 0.00017838369240052998
step: 30, loss: 0.001146501861512661
step: 40, loss: 0.00022626345162279904
step: 50, loss: 0.0004896338796243072
step: 60, loss: 0.004625467583537102
step: 70, loss: 0.010827296413481236
step: 80, loss: 0.0010023119393736124
step: 90, loss: 6.088630107115023e-05
step: 100, loss: 0.00022892907145433128
step: 110, loss: 0.004570616874843836
step: 120, loss: 7.25491699995473e-05
step: 130, loss: 0.00038083543768152595
step: 140, loss: 0.012219498865306377
step: 150, loss: 0.0003467446949798614
step: 160, loss: 0.033574748784303665
step: 170, loss: 0.0017785122618079185
step: 180, loss: 0.0025667904410511255
step: 190, loss: 5.691164915333502e-05
step: 200, loss: 0.00820330623537302
step: 210, loss: 0.020557070150971413
step: 220, loss: 0.0006471980013884604
step: 230, loss: 5.743445944972336e-05
step: 240, loss: 0.000541752262506634
step: 250, loss: 0.0004794571432285011

```

```

step: 260, loss: 0.013725070282816887
step: 270, loss: 7.977891073096544e-05
step: 280, loss: 0.026254473254084587
step: 290, loss: 0.0013546159025281668
step: 300, loss: 0.00011964637815253809
step: 310, loss: 0.05861320346593857
step: 320, loss: 0.0009133177809417248
step: 330, loss: 0.0016140304505825043
step: 340, loss: 0.009046472609043121
step: 350, loss: 0.0017401595832780004
step: 360, loss: 0.00017484014097135514
step: 370, loss: 0.0015358439413830638
step: 380, loss: 0.0026648547500371933
step: 390, loss: 0.0003574453585315496
step: 400, loss: 0.0013897984754294157
step: 410, loss: 0.013157380744814873
step: 420, loss: 0.00023040738597046584
step: 430, loss: 0.0002508683246560395
step: 440, loss: 0.00022029691899660975
step: 450, loss: 0.001618409063667059
step: 460, loss: 0.0002027719747275114
step: 470, loss: 0.0009160461486317217
step: 480, loss: 0.0001019308838294819
step: 490, loss: 0.00017919536912813783
step: 500, loss: 0.004765987861901522
step: 510, loss: 2.7318714273860678e-05
step: 520, loss: 0.001030607963912189
step: 530, loss: 0.00022451419499702752
/usr/local/lib/python3.7/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=0.9529953917050692, f1=0.9547553093259464,
best_f1=0.9532967032967032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.0068493288126774e-05
step: 10, loss: 8.573394734412432e-05
step: 20, loss: 6.82159443385899e-05
step: 30, loss: 0.005307900253683329
step: 40, loss: 0.00011862079554703087
step: 50, loss: 0.000157856396981515
step: 60, loss: 4.112156602786854e-05
step: 70, loss: 0.0002652684925124049
step: 80, loss: 0.002005157293751836
step: 90, loss: 0.024153703823685646
step: 100, loss: 0.00022978360357228667
step: 110, loss: 0.0001138238440034911

```

```
step: 120, loss: 0.00018001228454522789
step: 130, loss: 4.8212528781732544e-05
step: 140, loss: 3.568213287508115e-05
step: 150, loss: 0.00010589459270704538
step: 160, loss: 0.00011701000039465725
step: 170, loss: 0.0002624545886646956
step: 180, loss: 4.047978654853068e-05
step: 190, loss: 0.0011332659050822258
step: 200, loss: 3.244736581109464e-05
step: 210, loss: 0.0009040810400620103
step: 220, loss: 0.010305906645953655
step: 230, loss: 2.9181112040532753e-05
step: 240, loss: 0.004599784966558218
step: 250, loss: 0.00013472145656123757
step: 260, loss: 0.05461588874459267
step: 270, loss: 0.0006704546976834536
step: 280, loss: 0.0001890143903438002
step: 290, loss: 0.0001860127376858145
step: 300, loss: 0.0005093090003356338
step: 310, loss: 0.0004896847531199455
step: 320, loss: 0.0018580637406557798
step: 330, loss: 6.567913806065917e-05
step: 340, loss: 0.0012825302546843886
step: 350, loss: 8.334620360983536e-05
step: 360, loss: 0.00029396140598692
step: 370, loss: 0.0005948621546849608
step: 380, loss: 0.0023287904914468527
step: 390, loss: 0.001567250699736178
step: 400, loss: 0.000812023296020925
step: 410, loss: 0.0013288523769006133
step: 420, loss: 0.0019957772456109524
step: 430, loss: 0.002586517482995987
step: 440, loss: 4.44760407845024e-05
step: 450, loss: 0.0011384085519239306
step: 460, loss: 0.0014822555240243673
step: 470, loss: 0.0007831510156393051
step: 480, loss: 0.006048052571713924
step: 490, loss: 5.888538726139814e-05
step: 500, loss: 0.0006201117648743093
step: 510, loss: 0.001193137839436531
step: 520, loss: 0.00018472172087058425
step: 530, loss: 0.002783657982945442
/usr/local/lib/python3.7/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=0.951545916012921, f1=0.9530386740331491,
best_f1=0.9532967032967032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0034214735496789217
step: 10, loss: 0.00024206048692576587
step: 20, loss: 0.0001125814305851236
step: 30, loss: 6.980955367907882e-05
step: 40, loss: 0.004146753344684839
step: 50, loss: 0.0017087050946429372
step: 60, loss: 8.978504047263414e-05
step: 70, loss: 0.0002404950646450743
step: 80, loss: 0.023452162742614746
step: 90, loss: 0.005194354336708784
step: 100, loss: 0.0024356108624488115
step: 110, loss: 9.011432848637924e-05
step: 120, loss: 7.24150741007179e-05
step: 130, loss: 3.986513547715731e-05
step: 140, loss: 7.616660150233656e-05
step: 150, loss: 5.274203067529015e-05
step: 160, loss: 0.0016608929727226496
step: 170, loss: 2.7640629923553206e-05
step: 180, loss: 0.0003858281997963786
step: 190, loss: 9.743691043695435e-05
step: 200, loss: 0.0011392752639949322
step: 210, loss: 3.8077527278801426e-05
step: 220, loss: 3.7193618481978774e-05
step: 230, loss: 2.3605440219398588e-05
step: 240, loss: 0.001345172873698175
step: 250, loss: 9.303934348281473e-05
step: 260, loss: 4.036082100355998e-05
step: 270, loss: 0.00755577115342021
step: 280, loss: 0.03426589444279671
step: 290, loss: 0.0019289086339995265
step: 300, loss: 0.00039198281592689455
step: 310, loss: 0.07946666330099106
step: 320, loss: 0.0001462311192881316
step: 330, loss: 0.0021764696575701237
step: 340, loss: 0.0002394901675870642
step: 350, loss: 1.7187823686981574e-05
step: 360, loss: 0.006327333860099316
step: 370, loss: 0.0004212836211081594
step: 380, loss: 4.9786147428676486e-05
step: 390, loss: 0.001842128112912178
step: 400, loss: 3.3152580726891756e-05
step: 410, loss: 0.00012376676022540778
step: 420, loss: 0.001259199227206409
step: 430, loss: 0.00047586983419023454
step: 440, loss: 0.00038978474913164973
step: 450, loss: 0.0001254160306416452
```

```

step: 460, loss: 6.752913759555668e-05
step: 470, loss: 0.003630306338891387
step: 480, loss: 0.0002710028493311256
step: 490, loss: 3.569387990864925e-05
step: 500, loss: 2.7941026928601786e-05
step: 510, loss: 0.019946400076150894
step: 520, loss: 0.005088197998702526
step: 530, loss: 4.232683568261564e-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=0.9495718363463369, f1=0.9444444444444445,
best_f1=0.9532967032967032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.031567030120641e-05
step: 10, loss: 0.22551682591438293
step: 20, loss: 7.795335841365159e-05
step: 30, loss: 2.0510933609330095e-05
step: 40, loss: 0.0011510227341204882
step: 50, loss: 0.002948432695120573
step: 60, loss: 4.58327449450735e-05
step: 70, loss: 6.227928679436445e-05
step: 80, loss: 0.000766459503211081
step: 90, loss: 2.8310994821367785e-05
step: 100, loss: 0.0002443989797029644
step: 110, loss: 0.0002382005623076111
step: 120, loss: 3.969246245105751e-05
step: 130, loss: 2.095727904816158e-05
step: 140, loss: 5.280121331452392e-05
step: 150, loss: 0.0014718429883942008
step: 160, loss: 0.00048411579336971045
step: 170, loss: 0.0014755136799067259
step: 180, loss: 0.0008580776047892869
step: 190, loss: 0.00025951830320991576
step: 200, loss: 9.080535528482869e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 2.1613108401652426e-05
step: 220, loss: 4.2084153392352164e-05
step: 230, loss: 0.0037942638155072927
step: 240, loss: 0.0024073394015431404
step: 250, loss: 1.6815645722090267e-05
step: 260, loss: 3.806740278378129e-05
step: 270, loss: 0.0008578365086577833
step: 280, loss: 2.8458371161832474e-05
step: 290, loss: 1.2390036317810882e-05
step: 300, loss: 1.4010625818627886e-05

```



```

step: 310, loss: 8.322581561515108e-05
step: 320, loss: 2.6928033548756503e-05
step: 330, loss: 9.909489017445594e-05
step: 340, loss: 0.00041548401350155473
step: 350, loss: 1.6982601664494723e-05
step: 360, loss: 0.10201015323400497
step: 370, loss: 0.0005516550736501813
step: 380, loss: 3.078629742958583e-05
step: 390, loss: 0.00010096117330249399
step: 400, loss: 0.00011941533739445731
step: 410, loss: 3.072699109907262e-05
step: 420, loss: 0.00028506311355158687
step: 430, loss: 5.2805524319410324e-05
step: 440, loss: 9.90174066828331e-06
step: 450, loss: 0.001211012015119195
step: 460, loss: 0.0007372578256763518
step: 470, loss: 0.035070594400167465
step: 480, loss: 2.603082248242572e-05
step: 490, loss: 2.8949853003723547e-05
step: 500, loss: 0.0003605267556849867
step: 510, loss: 0.0002580881700851023
step: 520, loss: 3.253934846725315e-05
step: 530, loss: 2.6513698685448617e-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=0.9488151658767773, f1=0.9495238095238094,
best_f1=0.9532967032967032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030081547796726227
step: 10, loss: 0.0002087561006192118
step: 20, loss: 1.7943859347724356e-05
step: 30, loss: 0.00037762182182632387
step: 40, loss: 6.881842273287475e-05
step: 50, loss: 0.0003217355697415769
step: 60, loss: 4.53319626103621e-05
step: 70, loss: 4.241341957822442e-05
step: 80, loss: 0.004285247065126896
step: 90, loss: 1.639056245039683e-05
step: 100, loss: 3.087978620897047e-05
step: 110, loss: 0.00016806699568405747
step: 120, loss: 1.190208422485739e-05
step: 130, loss: 1.1522166460053995e-05
step: 140, loss: 9.096758731175214e-05
step: 150, loss: 0.00010847693192772567
step: 160, loss: 1.655847518122755e-05

```

```

step: 170, loss: 9.17485522227052e-05
step: 180, loss: 0.00011422267562011257
step: 190, loss: 4.176608854322694e-05
step: 200, loss: 1.2516797141870484e-05
step: 210, loss: 0.00015277510101441294
step: 220, loss: 3.248569555580616e-05
step: 230, loss: 0.00016213537310250103
step: 240, loss: 0.002038288163021207
step: 250, loss: 0.0017977439565584064
step: 260, loss: 0.00014324979565571994
step: 270, loss: 0.004896858241409063
step: 280, loss: 0.00021242332877591252
step: 290, loss: 1.7094838767661713e-05
step: 300, loss: 2.6898240321315825e-05
step: 310, loss: 2.4677619876456447e-05
step: 320, loss: 2.6369181796326302e-05
step: 330, loss: 2.8047286832588725e-05
step: 340, loss: 1.2658360901696142e-05
step: 350, loss: 1.0751048648671713e-05
step: 360, loss: 4.122857717447914e-05
step: 370, loss: 0.0009233345044776797
step: 380, loss: 0.00017497649241704494
step: 390, loss: 0.00050796550931409
step: 400, loss: 0.0007565468549728394
step: 410, loss: 1.1786441973526962e-05
step: 420, loss: 1.2643407899304293e-05
step: 430, loss: 6.573417340405285e-05
step: 440, loss: 0.0006159330951049924
step: 450, loss: 1.5280891602742486e-05
step: 460, loss: 0.03658638894557953
step: 470, loss: 9.365307960251812e-06
step: 480, loss: 0.00021682700025849044
step: 490, loss: 3.394378290977329e-05
step: 500, loss: 0.00039023967110551894
step: 510, loss: 0.025186873972415924
step: 520, loss: 0.0002040304971160367
step: 530, loss: 0.0033941329456865788
/usr/local/lib/python3.7/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=0.9538891476478809, f1=0.9541454377026402,
best_f1=0.9532967032967032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.1410405932110734e-05
step: 10, loss: 0.0005627868813462555
step: 20, loss: 0.001149081508629024

```

step: 30, loss: 0.003144456073641777
step: 40, loss: 1.192823128803866e-05
step: 50, loss: 0.0011380136711522937
step: 60, loss: 0.0005239741294644773
step: 70, loss: 1.31945680550416e-05
step: 80, loss: 1.121668810810661e-05
step: 90, loss: 1.795484422473237e-05
step: 100, loss: 6.638425020355498e-06
step: 110, loss: 0.00012850186612922698
step: 120, loss: 1.166742731584236e-05
step: 130, loss: 1.1272554729657713e-05
step: 140, loss: 1.3503850823326502e-05
step: 150, loss: 1.801798316591885e-05
step: 160, loss: 0.001582309021614492
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 1.0091659532918129e-05
step: 180, loss: 0.03253290802240372
step: 190, loss: 0.0031544289086014032
step: 200, loss: 0.0015173106221482158
step: 210, loss: 2.2942100258660503e-05
step: 220, loss: 0.01677348092198372
step: 230, loss: 0.0005071079358458519
step: 240, loss: 1.4111077689449303e-05
step: 250, loss: 1.5045608961372636e-05
step: 260, loss: 7.87148474046262e-06
step: 270, loss: 1.4990157978900243e-05
step: 280, loss: 7.1674071477900725e-06
step: 290, loss: 0.00017797671898733824
step: 300, loss: 1.0955931429634802e-05
step: 310, loss: 0.05381861329078674
step: 320, loss: 1.0896364074142184e-05
step: 330, loss: 1.260996941709891e-05
step: 340, loss: 2.0614987079170533e-05
step: 350, loss: 1.2054786566295661e-05
step: 360, loss: 1.1346984138072003e-05
step: 370, loss: 0.00046716819633729756
step: 380, loss: 1.1767958312702831e-05
step: 390, loss: 1.6871334082679823e-05
step: 400, loss: 0.0007126607233658433
step: 410, loss: 4.795116183231585e-05
step: 420, loss: 1.0907482646871358e-05
step: 430, loss: 7.010920398897724e-06
step: 440, loss: 4.223778523737565e-05
step: 450, loss: 0.004248516634106636
step: 460, loss: 6.481935997726396e-05
step: 470, loss: 9.6148714874289e-06
step: 480, loss: 0.0022115036845207214
step: 490, loss: 6.048839713912457e-05

```

step: 500, loss: 0.0005536886164918542
step: 510, loss: 1.319462444371311e-05
step: 520, loss: 1.2583835086843465e-05
step: 530, loss: 6.5862805058714e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9531322505800465, f1=0.9542302357836338,
best_f1=0.9532967032967032
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:35, 163.87it/s]
load_f1 = 0.9552376557452701
real_f1 = 0.9533487297921477
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 137.72it/s]

```

1.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled          : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5082610845565796
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41014501452445984
step: 20, loss: 0.41079485416412354
step: 30, loss: 0.34032168984413147
step: 40, loss: 0.3088364601135254
step: 50, loss: 0.46108531951904297
step: 60, loss: 0.4913254678249359
step: 70, loss: 0.30852770805358887
step: 80, loss: 0.37282803654670715
step: 90, loss: 0.29835572838783264
step: 100, loss: 0.2358703762292862
step: 110, loss: 0.2614438533782959
step: 120, loss: 0.37002983689308167
step: 130, loss: 0.242886483669281
step: 140, loss: 0.469072550535202
step: 150, loss: 0.39382943511009216
step: 160, loss: 0.556541919708252
step: 170, loss: 0.2594667375087738
step: 180, loss: 0.36675500869750977
step: 190, loss: 0.6922257542610168
step: 200, loss: 0.3952532708644867
step: 210, loss: 0.5115228891372681
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3766460716724396
step: 10, loss: 0.2036057412624359
step: 20, loss: 0.4983813762664795
step: 30, loss: 0.4959862232208252
step: 40, loss: 0.47283560037612915
step: 50, loss: 0.2298993468284607
step: 60, loss: 0.307115375995636
step: 70, loss: 0.47116324305534363
step: 80, loss: 0.31518784165382385
step: 90, loss: 0.37581735849380493
step: 100, loss: 0.5014696717262268
step: 110, loss: 0.38842788338661194
step: 120, loss: 0.23501017689704895
step: 130, loss: 0.166880264878273
step: 140, loss: 0.24161306023597717
step: 150, loss: 0.4485141932964325
step: 160, loss: 0.16808566451072693
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.6293199062347412
step: 180, loss: 0.30978530645370483
step: 190, loss: 0.3082669973373413
step: 200, loss: 0.15426571667194366
step: 210, loss: 0.3066663146018982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2541778087615967
step: 10, loss: 0.24339091777801514
step: 20, loss: 0.4812225103378296
step: 30, loss: 0.23416616022586823
step: 40, loss: 0.45059701800346375
step: 50, loss: 0.4598439931869507
step: 60, loss: 0.48858317732810974
step: 70, loss: 0.18350525200366974

```

```

step: 80, loss: 0.469148725271225
step: 90, loss: 0.23043225705623627
step: 100, loss: 0.36992132663726807
step: 110, loss: 0.23372003436088562
step: 120, loss: 0.25929203629493713
step: 130, loss: 0.15833240747451782
step: 140, loss: 0.34718555212020874
step: 150, loss: 0.32015788555145264
step: 160, loss: 0.2288951426744461
step: 170, loss: 0.3722613453865051
step: 180, loss: 0.2512328028678894
step: 190, loss: 0.167636439204216
step: 200, loss: 0.25133809447288513
step: 210, loss: 0.24972401559352875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30829378962516785
step: 10, loss: 0.30165228247642517
step: 20, loss: 0.30636751651763916
step: 30, loss: 0.2343539446592331
step: 40, loss: 0.23516327142715454
step: 50, loss: 0.2404763549566269
step: 60, loss: 0.4950231611728668
step: 70, loss: 0.25226229429244995
step: 80, loss: 0.23952162265777588
step: 90, loss: 0.4674607217311859
step: 100, loss: 0.38004085421562195
step: 110, loss: 0.6038954257965088
step: 120, loss: 0.3750755488872528
step: 130, loss: 0.6812676191329956
step: 140, loss: 0.49355781078338623
step: 150, loss: 0.3791744112968445
step: 160, loss: 0.306303471326828
step: 170, loss: 0.17750364542007446
step: 180, loss: 0.08599672466516495
step: 190, loss: 0.1716042160987854
step: 200, loss: 0.32651910185813904
step: 210, loss: 0.3785044252872467
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,

```



```

best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3904838562011719
step: 10, loss: 0.3164185583591461
step: 20, loss: 0.3071117103099823
step: 30, loss: 0.23857487738132477
step: 40, loss: 0.44951921701431274
step: 50, loss: 0.37788131833076477
step: 60, loss: 0.3884945511817932
step: 70, loss: 0.23828230798244476
step: 80, loss: 0.44211313128471375
step: 90, loss: 0.43223217129707336
step: 100, loss: 0.24264521896839142
step: 110, loss: 0.1652337908744812
step: 120, loss: 0.23585985600948334
step: 130, loss: 0.31361523270606995
step: 140, loss: 0.5241647958755493
step: 150, loss: 0.31911709904670715
step: 160, loss: 0.2494027465581894
step: 170, loss: 0.375164270401001
step: 180, loss: 0.23539315164089203
step: 190, loss: 0.5201518535614014
step: 200, loss: 0.37371140718460083
step: 210, loss: 0.2423170953989029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1683349311351776
step: 10, loss: 0.3757074475288391
step: 20, loss: 0.3099704682826996
step: 30, loss: 0.2392355352640152
step: 40, loss: 0.32367056608200073
step: 50, loss: 0.5996413826942444
step: 60, loss: 0.3090117573738098
step: 70, loss: 0.45106586813926697
step: 80, loss: 0.3131057322025299
step: 90, loss: 0.31090134382247925
step: 100, loss: 0.31053629517555237
step: 110, loss: 0.3834516704082489
step: 120, loss: 0.2539282739162445
step: 130, loss: 0.2401513159275055

```

```

step: 140, loss: 0.38092419505119324
step: 150, loss: 0.19356058537960052
step: 160, loss: 0.16681864857673645
step: 170, loss: 0.4529056251049042
step: 180, loss: 0.38564735651016235
step: 190, loss: 0.4572487771511078
step: 200, loss: 0.3135196268558502
step: 210, loss: 0.3739207684993744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.46151456236839294
step: 10, loss: 0.3248922526836395
step: 20, loss: 0.44124293327331543
step: 30, loss: 0.24641208350658417
step: 40, loss: 0.24701803922653198
step: 50, loss: 0.3731505572795868
step: 60, loss: 0.3142114579677582
step: 70, loss: 0.2512874901294708
step: 80, loss: 0.5145779848098755
step: 90, loss: 0.38061434030532837
step: 100, loss: 0.4424697458744049
step: 110, loss: 0.3068847358226776
step: 120, loss: 0.30863314867019653
step: 130, loss: 0.4671308696269989
step: 140, loss: 0.3143220841884613
step: 150, loss: 0.30753737688064575
step: 160, loss: 0.550458550453186
step: 170, loss: 0.6056978106498718
step: 180, loss: 0.2428511530160904
step: 190, loss: 0.2443663626909256
step: 200, loss: 0.31069374084472656
step: 210, loss: 0.3132762908935547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6051820516586304
step: 10, loss: 0.31010302901268005

```

```

step: 20, loss: 0.2598622739315033
step: 30, loss: 0.254655659198761
step: 40, loss: 0.25337085127830505
step: 50, loss: 0.17495065927505493
step: 60, loss: 0.1897534430027008
step: 70, loss: 0.47982820868492126
step: 80, loss: 0.3050633668899536
step: 90, loss: 0.37442338466644287
step: 100, loss: 0.6068107485771179
step: 110, loss: 0.3734467625617981
step: 120, loss: 0.3189573884010315
step: 130, loss: 0.1686411201953888
step: 140, loss: 0.3146441876888275
step: 150, loss: 0.4679372310638428
step: 160, loss: 0.4530183970928192
step: 170, loss: 0.5150007009506226
step: 180, loss: 0.3105480968952179
step: 190, loss: 0.24273933470249176
step: 200, loss: 0.4594360291957855
step: 210, loss: 0.382914662361145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5756051540374756
step: 10, loss: 0.2961004078388214
step: 20, loss: 0.3784573972225189
step: 30, loss: 0.16923809051513672
step: 40, loss: 0.16293275356292725
step: 50, loss: 0.4660361707210541
step: 60, loss: 0.17642930150032043
step: 70, loss: 0.3891627788543701
step: 80, loss: 0.31580013036727905
step: 90, loss: 0.3083171844482422
step: 100, loss: 0.46228259801864624
step: 110, loss: 0.5322257876396179
step: 120, loss: 0.3763021230697632
step: 130, loss: 0.24263499677181244
step: 140, loss: 0.6034188866615295
step: 150, loss: 0.11110921949148178
step: 160, loss: 0.3781305253505707
step: 170, loss: 0.3030232787132263
step: 180, loss: 0.4737456440925598
step: 190, loss: 0.3175540268421173

```

```

step: 200, loss: 0.3077542781829834
step: 210, loss: 0.3099668025970459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.22597778850796718, f1=0.2816553428042001,
best_f1=0.2816553428042001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.351164847612381
step: 10, loss: 0.381942480802536
step: 20, loss: 0.24550005793571472
step: 30, loss: 0.1550455391407013
step: 40, loss: 0.25981447100639343
step: 50, loss: 0.31426572799682617
step: 60, loss: 0.13435699045658112
step: 70, loss: 0.29874086380004883
step: 80, loss: 0.13243886828422546
step: 90, loss: 0.28290486335754395
step: 100, loss: 0.3187835216522217
step: 110, loss: 0.15061943233013153
step: 120, loss: 0.4355504512786865
step: 130, loss: 0.12633110582828522
step: 140, loss: 0.19300805032253265
step: 150, loss: 0.2114030420780182
step: 160, loss: 0.25058045983314514
step: 170, loss: 0.1313183754682541
step: 180, loss: 0.18006743490695953
step: 190, loss: 0.1466551423072815
step: 200, loss: 0.3535596430301666
step: 210, loss: 0.1289663314819336
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6515151515151516, f1=0.64, best_f1=0.64
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19307731091976166

```

```

step: 10, loss: 0.20826908946037292
step: 20, loss: 0.11682132631540298
step: 30, loss: 0.055593784898519516
step: 40, loss: 0.2365250289440155
step: 50, loss: 0.46788302063941956
step: 60, loss: 0.16101717948913574
step: 70, loss: 0.14501169323921204
step: 80, loss: 0.19457583129405975
step: 90, loss: 0.27489715814590454
step: 100, loss: 0.21695536375045776
step: 110, loss: 0.24927492439746857
step: 120, loss: 0.15846456587314606
step: 130, loss: 0.07756292074918747
step: 140, loss: 0.11391861736774445
step: 150, loss: 0.26853612065315247
step: 160, loss: 0.16390679776668549
step: 170, loss: 0.13620741665363312
step: 180, loss: 0.11256135255098343
step: 190, loss: 0.398150771856308
step: 200, loss: 0.052947964519262314
step: 210, loss: 0.25491780042648315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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 11: dev_f1=0.6770601336302895, f1=0.6776470588235295,
best_f1=0.6776470588235295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14702048897743225
step: 10, loss: 0.16935555636882782
step: 20, loss: 0.29656457901000977
step: 30, loss: 0.1769593358039856
step: 40, loss: 0.055412229150533676
step: 50, loss: 0.2083292156457901
step: 60, loss: 0.040874235332012177
step: 70, loss: 0.205942302942276
step: 80, loss: 0.123940110206604
step: 90, loss: 0.18143656849861145
step: 100, loss: 0.06122102215886116
step: 110, loss: 0.24720287322998047
step: 120, loss: 0.04256405681371689
step: 130, loss: 0.26007771492004395
step: 140, loss: 0.10919976234436035

```

```

step: 150, loss: 0.01680767349898815
step: 160, loss: 0.12511275708675385
step: 170, loss: 0.12266163527965546
step: 180, loss: 0.19064557552337646
step: 190, loss: 0.12785100936889648
step: 200, loss: 0.07009799778461456
step: 210, loss: 0.08088894933462143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6851851851851851, f1=0.6893939393939394,
best_f1=0.6893939393939394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13855360448360443
step: 10, loss: 0.05953998491168022
step: 20, loss: 0.07486278563737869
step: 30, loss: 0.07188326120376587
step: 40, loss: 0.1405216008424759
step: 50, loss: 0.21925105154514313
step: 60, loss: 0.22647924721240997
step: 70, loss: 0.12432003021240234
step: 80, loss: 0.136163130402565
step: 90, loss: 0.13873668015003204
step: 100, loss: 0.2315378040075302
step: 110, loss: 0.13252314925193787
step: 120, loss: 0.20078523457050323
step: 130, loss: 0.10363569110631943
step: 140, loss: 0.1617993265390396
step: 150, loss: 0.0705592930316925
step: 160, loss: 0.11607656627893448
step: 170, loss: 0.12234163284301758
step: 180, loss: 0.07709937542676926
step: 190, loss: 0.06858418881893158
step: 200, loss: 0.11844804137945175
step: 210, loss: 0.2002187818288803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6909090909090909, f1=0.7193347193347193,
best_f1=0.7193347193347193
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08152594417333603
step: 10, loss: 0.11539901793003082
step: 20, loss: 0.3017851412296295
step: 30, loss: 0.08060431480407715
step: 40, loss: 0.13703522086143494
step: 50, loss: 0.1645398736000061
step: 60, loss: 0.1231124997138977
step: 70, loss: 0.10446176677942276
step: 80, loss: 0.1530875861644745
step: 90, loss: 0.021807216107845306
step: 100, loss: 0.051577288657426834
step: 110, loss: 0.13971391320228577
step: 120, loss: 0.05524824559688568
step: 130, loss: 0.13729292154312134
step: 140, loss: 0.09429701417684555
step: 150, loss: 0.07133185863494873
step: 160, loss: 0.01752004772424698
step: 170, loss: 0.060399819165468216
step: 180, loss: 0.09173555672168732
step: 190, loss: 0.0639820396900177
step: 200, loss: 0.04262080788612366
step: 210, loss: 0.10698921233415604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6973180076628351, f1=0.7151277013752456,
best_f1=0.7151277013752456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.139412060379982
step: 10, loss: 0.10755319893360138
step: 20, loss: 0.10636653006076813
step: 30, loss: 0.1354776918888092
step: 40, loss: 0.07723593711853027
step: 50, loss: 0.14536848664283752
step: 60, loss: 0.18181759119033813
step: 70, loss: 0.09151805192232132
step: 80, loss: 0.035752762109041214

```

```

step: 90, loss: 0.04472712054848671
step: 100, loss: 0.06967778503894806
step: 110, loss: 0.21804925799369812
step: 120, loss: 0.19006860256195068
step: 130, loss: 0.06510665267705917
step: 140, loss: 0.09661346673965454
step: 150, loss: 0.20495489239692688
step: 160, loss: 0.27696332335472107
step: 170, loss: 0.06311351805925369
step: 180, loss: 0.051372841000556946
step: 190, loss: 0.04197785258293152
step: 200, loss: 0.04841308668255806
step: 210, loss: 0.1498918980360031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.7034220532319392, f1=0.7123287671232877,
best_f1=0.7123287671232877
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 [00:09, 234.19it/s]
load_f1 = 0.6984732824427482
real_f1 = 0.6895238095238096
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 138.63it/s]

```

1.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

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

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

/usr/local/lib/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: 0.4252094626426697

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

step: 10, loss: 0.5074025392532349

step: 20, loss: 0.3230299651622772

step: 30, loss: 0.39606431126594543

step: 40, loss: 0.2618615925312042

step: 50, loss: 0.3195364475250244

step: 60, loss: 0.47071728110313416

step: 70, loss: 0.413546621799469

step: 80, loss: 0.16426683962345123

step: 90, loss: 0.30158084630966187

step: 100, loss: 0.4328615367412567

step: 110, loss: 0.2507537603378296

```

step: 120, loss: 0.34418392181396484
step: 130, loss: 0.3042128384113312
step: 140, loss: 0.17451795935630798
step: 150, loss: 0.2945065200328827
step: 160, loss: 0.231487438082695
step: 170, loss: 0.33662137389183044
step: 180, loss: 0.196697399020195
step: 190, loss: 0.14246664941310883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4773413897280967, f1=0.49132947976878616,
best_f1=0.49132947976878616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27278932929039
step: 10, loss: 0.21860387921333313
step: 20, loss: 0.43023210763931274
step: 30, loss: 0.15768837928771973
step: 40, loss: 0.17699164152145386
step: 50, loss: 0.40356045961380005
step: 60, loss: 0.397851824760437
step: 70, loss: 0.08039404451847076
step: 80, loss: 0.06190670281648636
step: 90, loss: 0.1528944969177246
step: 100, loss: 0.20612181723117828
step: 110, loss: 0.048176735639572144
step: 120, loss: 0.08564412593841553
step: 130, loss: 0.09515272825956345
step: 140, loss: 0.19008322060108185
step: 150, loss: 0.292635977268219
step: 160, loss: 0.08582141250371933
step: 170, loss: 0.013442880474030972
step: 180, loss: 0.06912176311016083
step: 190, loss: 0.020735537633299828
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8125, f1=0.8167539267015705, best_f1=0.8167539267015705

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10185851901769638
step: 10, loss: 0.10436426103115082
step: 20, loss: 0.11736597865819931
step: 30, loss: 0.06474632769823074
step: 40, loss: 0.028421759605407715
step: 50, loss: 0.1393248438835144
step: 60, loss: 0.1284089833498001
step: 70, loss: 0.1276063174009323
step: 80, loss: 0.1607537418603897
step: 90, loss: 0.11273709684610367
step: 100, loss: 0.08521458506584167
step: 110, loss: 0.38788536190986633
step: 120, loss: 0.23531609773635864
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.07407566905021667
step: 140, loss: 0.08872903138399124
step: 150, loss: 0.1525009274482727
step: 160, loss: 0.0894496738910675
step: 170, loss: 0.32635530829429626
step: 180, loss: 0.3407549560070038
step: 190, loss: 0.12127196043729782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8290155440414507, f1=0.8373333333333334,
best_f1=0.8373333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03598654270172119
step: 10, loss: 0.09476517140865326
step: 20, loss: 0.01533539965748787
step: 30, loss: 0.007293581031262875
step: 40, loss: 0.0365324467420578
step: 50, loss: 0.10822143405675888
step: 60, loss: 0.1456604301929474
step: 70, loss: 0.02276284247636795
step: 80, loss: 0.015214655548334122
step: 90, loss: 0.024441787973046303
step: 100, loss: 0.021319007501006126
step: 110, loss: 0.1182490810751915

```

```

step: 120, loss: 0.02975689433515072
step: 130, loss: 0.01936749741435051
step: 140, loss: 0.17964798212051392
step: 150, loss: 0.057544928044080734
step: 160, loss: 0.0178458821028471
step: 170, loss: 0.05045991763472557
step: 180, loss: 0.03741935268044472
step: 190, loss: 0.011541126295924187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8409703504043127, f1=0.8435013262599469,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13365095853805542
step: 10, loss: 0.030147001147270203
step: 20, loss: 0.007504328154027462
step: 30, loss: 0.052870240062475204
step: 40, loss: 0.0062552583403885365
step: 50, loss: 0.008789845742285252
step: 60, loss: 0.032033443450927734
step: 70, loss: 0.036843229085206985
step: 80, loss: 0.009906049817800522
step: 90, loss: 0.06344093382358551
step: 100, loss: 0.03815256059169769
step: 110, loss: 0.17039915919303894
step: 120, loss: 0.08857668191194534
step: 130, loss: 0.2900068759918213
step: 140, loss: 0.028978731483221054
step: 150, loss: 0.029581686481833458
step: 160, loss: 0.008432740345597267
step: 170, loss: 0.013647847808897495
step: 180, loss: 0.02698453888297081
step: 190, loss: 0.05724861100316048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8503937007874015, f1=0.8435013262599469,

```

```

best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012633463367819786
step: 10, loss: 0.013562025502324104
step: 20, loss: 0.008920871652662754
step: 30, loss: 0.025695674121379852
step: 40, loss: 0.049520812928676605
step: 50, loss: 0.006710301619023085
step: 60, loss: 0.007715859916061163
step: 70, loss: 0.011486150324344635
step: 80, loss: 0.022138528525829315
step: 90, loss: 0.03456222265958786
step: 100, loss: 0.12351547181606293
step: 110, loss: 0.0028372756205499172
step: 120, loss: 0.0460817851126194
step: 130, loss: 0.06439484655857086
step: 140, loss: 0.05074948072433472
step: 150, loss: 0.008299188688397408
step: 160, loss: 0.013123472221195698
step: 170, loss: 0.3707350790500641
step: 180, loss: 0.07209734618663788
step: 190, loss: 0.01466903556138277
/usr/local/lib/python3.7/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=0.8429752066115702, f1=0.8412256267409471,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004169621504843235
step: 10, loss: 0.0035439501516520977
step: 20, loss: 0.0030170450918376446
step: 30, loss: 0.010395531542599201
step: 40, loss: 0.007005246821790934
step: 50, loss: 0.0036126987542957067
step: 60, loss: 0.0409012995660305
step: 70, loss: 0.0026514416094869375
step: 80, loss: 0.0025326129980385303
step: 90, loss: 0.02671797387301922
step: 100, loss: 0.1574547439813614
step: 110, loss: 0.005181537009775639
step: 120, loss: 0.037109192460775375
step: 130, loss: 0.08570721745491028
step: 140, loss: 0.040938615798950195
step: 150, loss: 0.05906103923916817

```

```

step: 160, loss: 0.2141634076833725
step: 170, loss: 0.16709692776203156
step: 180, loss: 0.0175014715641737
step: 190, loss: 0.028941353783011436
/usr/local/lib/python3.7/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=0.8314606741573033, f1=0.8444444444444443,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019523410126566887
step: 10, loss: 0.01855810545384884
step: 20, loss: 0.013874666765332222
step: 30, loss: 0.019352877512574196
step: 40, loss: 0.13554434478282928
step: 50, loss: 0.010289731435477734
step: 60, loss: 0.0073078167624771595
step: 70, loss: 0.0014686756767332554
step: 80, loss: 0.024355994537472725
step: 90, loss: 0.01711278222501278
step: 100, loss: 0.008322888053953648
step: 110, loss: 0.0547560453414917
step: 120, loss: 0.013408157974481583
step: 130, loss: 0.003513638162985444
step: 140, loss: 0.00964463409036398
step: 150, loss: 0.001732986536808312
step: 160, loss: 0.011767883785068989
step: 170, loss: 0.004589461255818605
step: 180, loss: 0.09085594117641449
step: 190, loss: 0.002854578197002411
/usr/local/lib/python3.7/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=0.83289124668435, f1=0.8364611260053619,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008548801764845848
step: 10, loss: 0.0012944808695465326
step: 20, loss: 0.04322657734155655
step: 30, loss: 0.0015926765045151114
step: 40, loss: 0.1397431194782257
step: 50, loss: 0.006284215487539768
step: 60, loss: 0.09820350259542465
step: 70, loss: 0.003507375717163086

```

```

step: 80, loss: 0.1156996414065361
step: 90, loss: 0.005357244051992893
step: 100, loss: 0.003434687154367566
step: 110, loss: 0.22340607643127441
step: 120, loss: 0.011073538102209568
step: 130, loss: 0.0058488138020038605
step: 140, loss: 0.15045399963855743
step: 150, loss: 0.0007997103966772556
step: 160, loss: 0.003589760512113571
step: 170, loss: 0.011303937062621117
step: 180, loss: 0.007041095290333033
step: 190, loss: 0.001157151535153389
/usr/local/lib/python3.7/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=0.8412256267409471, f1=0.8412256267409471,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008035743376240134
step: 10, loss: 0.000942840299103409
step: 20, loss: 0.0020391889847815037
step: 30, loss: 0.10636557638645172
step: 40, loss: 0.003971171099692583
step: 50, loss: 0.15970082581043243
step: 60, loss: 0.020527301356196404
step: 70, loss: 0.05681416764855385
step: 80, loss: 0.002828879514709115
step: 90, loss: 0.0006527628866024315
step: 100, loss: 0.03613140061497688
step: 110, loss: 0.07920941710472107
step: 120, loss: 0.006256532855331898
step: 130, loss: 0.0016778371063992381
step: 140, loss: 0.0011884734267368913
step: 150, loss: 0.12826530635356903
step: 160, loss: 0.10748225450515747
step: 170, loss: 0.0021693622693419456
step: 180, loss: 0.03202236816287041
step: 190, loss: 0.004479712340980768
/usr/local/lib/python3.7/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=0.8432432432432432, f1=0.8525469168900803,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.001593693858012557
step: 10, loss: 0.001267111860215664
step: 20, loss: 0.0973893254995346
step: 30, loss: 0.020176896825432777
step: 40, loss: 0.001175709068775177
step: 50, loss: 0.0016924610827118158
step: 60, loss: 0.0007353745750151575
step: 70, loss: 0.0010164682753384113
step: 80, loss: 0.00577140599489212
step: 90, loss: 0.009382251650094986
step: 100, loss: 0.0022784979082643986
step: 110, loss: 0.036565881222486496
step: 120, loss: 0.000994706992059946
step: 130, loss: 0.015966175124049187
step: 140, loss: 0.001988990232348442
step: 150, loss: 0.0046782284043729305
step: 160, loss: 0.003353371750563383
step: 170, loss: 0.003034257097169757
step: 180, loss: 0.018353145569562912
step: 190, loss: 0.00555774662643671
/usr/local/lib/python3.7/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=0.8463611859838275, f1=0.8469945355191256,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030475009232759476
step: 10, loss: 0.0024270927533507347
step: 20, loss: 0.058719802647829056
step: 30, loss: 0.007960342802107334
step: 40, loss: 0.0012907544150948524
step: 50, loss: 0.020210858434438705
step: 60, loss: 0.01865106076002121
step: 70, loss: 0.0071298484690487385
step: 80, loss: 0.0036934141535311937
step: 90, loss: 0.004567127209156752
step: 100, loss: 0.013851313851773739
step: 110, loss: 0.0007168222800828516
step: 120, loss: 0.0022722952999174595
step: 130, loss: 0.007125867530703545
step: 140, loss: 0.0007303732563741505
step: 150, loss: 0.016660576686263084
step: 160, loss: 0.0004304450994823128
step: 170, loss: 0.0011629186337813735
step: 180, loss: 0.0007298968848772347
step: 190, loss: 0.0006228974089026451

```

```

/usr/local/lib/python3.7/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=0.837772397094431, f1=0.8325123152709359,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01695023849606514
step: 10, loss: 0.001617110799998045
step: 20, loss: 0.00043749838368967175
step: 30, loss: 0.0071540214121341705
step: 40, loss: 0.0007878228207118809
step: 50, loss: 0.0008321556961163878
step: 60, loss: 0.001653266022913158
step: 70, loss: 0.011264222674071789
step: 80, loss: 0.006188609637320042
step: 90, loss: 0.004468906670808792
step: 100, loss: 0.0005590753862634301
step: 110, loss: 0.20846860110759735
step: 120, loss: 0.004626182373613119
step: 130, loss: 0.001782975741662085
step: 140, loss: 0.008234566077589989
step: 150, loss: 0.007443190552294254
step: 160, loss: 0.0009460397413931787
step: 170, loss: 0.0017910751048475504
step: 180, loss: 0.0025261654518544674
step: 190, loss: 0.02021808736026287
/usr/local/lib/python3.7/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=0.847645429362881, f1=0.8523676880222841,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03229374811053276
step: 10, loss: 0.006723261438310146
step: 20, loss: 0.003136663231998682
step: 30, loss: 0.0015901261940598488
step: 40, loss: 0.0011281727347522974
step: 50, loss: 0.007669389713555574
step: 60, loss: 0.0004310845397412777
step: 70, loss: 0.007086784113198519
step: 80, loss: 0.000501374714076519
step: 90, loss: 0.0005804107640869915
step: 100, loss: 0.0005863552214577794
step: 110, loss: 0.001139769796282053

```

```

step: 120, loss: 0.058690570294857025
step: 130, loss: 0.0014512748457491398
step: 140, loss: 0.0037531298585236073
step: 150, loss: 0.0010619205422699451
step: 160, loss: 0.0006848113262094557
step: 170, loss: 0.0007944187382236123
step: 180, loss: 0.04547611624002457
step: 190, loss: 0.007039690390229225
/usr/local/lib/python3.7/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=0.8488063660477454, f1=0.8693333333333334,
best_f1=0.8435013262599469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006415741518139839
step: 10, loss: 0.001041775569319725
step: 20, loss: 0.03229357674717903
step: 30, loss: 0.0005574235692620277
step: 40, loss: 0.10777389258146286
step: 50, loss: 0.0016079837223514915
step: 60, loss: 0.00111958768684417
step: 70, loss: 0.0020110916811972857
step: 80, loss: 0.006847884505987167
step: 90, loss: 0.00812560599297285
step: 100, loss: 0.0004088518617209047
step: 110, loss: 0.03252663463354111
step: 120, loss: 0.0012673686724156141
step: 130, loss: 0.0008109759073704481
step: 140, loss: 0.0012013016967102885
step: 150, loss: 0.0015884976601228118
step: 160, loss: 0.0012431311188265681
step: 170, loss: 0.006932099349796772
step: 180, loss: 0.0155815239995718
step: 190, loss: 0.03169535472989082
/usr/local/lib/python3.7/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=0.8503937007874015, f1=0.8713910761154856,
best_f1=0.8435013262599469
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:13, 150.53it/s]
load_f1 = 0.8061224489795918
real_f1 = 0.7764127764127764
733it [00:00, 3327.12it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 133.67it/s]

```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

```

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

```

```

--lm roberta \
--fp16 \
--save_model \
--da entry_swap \
--da span_shuffle \
--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.49926069378852844
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4756096303462982
step: 20, loss: 0.347630113363266
step: 30, loss: 0.47402727603912354
step: 40, loss: 0.49018940329551697
step: 50, loss: 0.31942829489707947
step: 60, loss: 0.5535739660263062
step: 70, loss: 0.3053838610649109
step: 80, loss: 0.23174940049648285
step: 90, loss: 0.2313050776720047
step: 100, loss: 0.1327790468931198
step: 110, loss: 0.3946821093559265
step: 120, loss: 0.3027650713920593
step: 130, loss: 0.30798810720443726
step: 140, loss: 0.3657178580760956
step: 150, loss: 0.33632344007492065
step: 160, loss: 0.36365044116973877
step: 170, loss: 0.28018680214881897
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2883333333333333, f1=0.3272058823529412,
best_f1=0.3272058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30708634853363037
step: 10, loss: 0.3440445065498352
step: 20, loss: 0.2275790423154831
step: 30, loss: 0.24108725786209106
step: 40, loss: 0.033877041190862656
step: 50, loss: 0.18434256315231323
step: 60, loss: 0.23813731968402863
step: 70, loss: 0.2275943011045456

```

```

step: 80, loss: 0.20377005636692047
step: 90, loss: 0.06283397227525711
step: 100, loss: 0.2813866138458252
step: 110, loss: 0.2002313882112503
step: 120, loss: 0.09983202815055847
step: 130, loss: 0.2866584062576294
step: 140, loss: 0.23246842622756958
step: 150, loss: 0.09602656215429306
step: 160, loss: 0.18472698330879211
step: 170, loss: 0.1087740883231163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7405660377358491, f1=0.7910112359550562,
best_f1=0.7910112359550562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26351621747016907
step: 10, loss: 0.04919417202472687
step: 20, loss: 0.021912453696131706
step: 30, loss: 0.12992152571678162
step: 40, loss: 0.03517081215977669
step: 50, loss: 0.13263142108917236
step: 60, loss: 0.06021507829427719
step: 70, loss: 0.15045106410980225
step: 80, loss: 0.08564194291830063
step: 90, loss: 0.4265533983707428
step: 100, loss: 0.06101543828845024
step: 110, loss: 0.022755175828933716
step: 120, loss: 0.03803986310958862
step: 130, loss: 0.09904167801141739
step: 140, loss: 0.21262226998806
step: 150, loss: 0.06044546887278557
step: 160, loss: 0.08732636272907257
step: 170, loss: 0.07715406268835068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8071065989847717, 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: 0.09652753919363022
step: 10, loss: 0.06280162185430527
step: 20, loss: 0.004275098908692598
step: 30, loss: 0.12197096645832062
step: 40, loss: 0.03315351903438568
step: 50, loss: 0.0322314128279686
step: 60, loss: 0.07832900434732437
step: 70, loss: 0.004580034874379635
step: 80, loss: 0.30439621210098267
step: 90, loss: 0.15539002418518066
step: 100, loss: 0.07640886306762695
step: 110, loss: 0.1820768564939499
step: 120, loss: 0.24976423382759094
step: 130, loss: 0.11167669296264648
step: 140, loss: 0.022021807730197906
step: 150, loss: 0.203548401594162
step: 160, loss: 0.010992404073476791
step: 170, loss: 0.11981622874736786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8353808353808354, f1=0.8717948717948718,
best_f1=0.8717948717948718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.074739970266819
step: 10, loss: 0.05597514286637306
step: 20, loss: 0.018657157197594643
step: 30, loss: 0.01849561557173729
step: 40, loss: 0.019493741914629936
step: 50, loss: 0.013813113793730736
step: 60, loss: 0.04961225390434265
step: 70, loss: 0.01899242214858532
step: 80, loss: 0.0018514448311179876
step: 90, loss: 0.02703562192618847
step: 100, loss: 0.02061532251536846
step: 110, loss: 0.03787851706147194
step: 120, loss: 0.10851456969976425
step: 130, loss: 0.025601113215088844

```



```

step: 140, loss: 0.061164431273937225
step: 150, loss: 0.05133449658751488
step: 160, loss: 0.009362460114061832
step: 170, loss: 0.005358167923986912
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8439897698209718, f1=0.8888888888888889,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12817487120628357
step: 10, loss: 0.008000175468623638
step: 20, loss: 0.12260451167821884
step: 30, loss: 0.0037489261012524366
step: 40, loss: 0.004671760369092226
step: 50, loss: 0.011580384336411953
step: 60, loss: 0.03274250775575638
step: 70, loss: 0.019885145127773285
step: 80, loss: 0.16282206773757935
step: 90, loss: 0.1656085103750229
step: 100, loss: 0.018498562276363373
step: 110, loss: 0.0626349076628685
step: 120, loss: 0.11225677281618118
step: 130, loss: 0.009430286474525928
step: 140, loss: 0.027920076623558998
step: 150, loss: 0.007159445434808731
step: 160, loss: 0.03296057879924774
step: 170, loss: 0.011484429240226746
/usr/local/lib/python3.7/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=0.8321167883211679, f1=0.8720379146919431,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03830839321017265
step: 10, loss: 0.003632865846157074
step: 20, loss: 0.0017514580395072699
step: 30, loss: 0.0017876580823212862
step: 40, loss: 0.0009967583464458585
step: 50, loss: 0.0023168521001935005

```

```

step: 60, loss: 0.004976989235728979
step: 70, loss: 0.0093537587672472
step: 80, loss: 0.032603975385427475
step: 90, loss: 0.019449274986982346
step: 100, loss: 0.002068551955744624
step: 110, loss: 0.0667153000831604
step: 120, loss: 0.01670689694583416
step: 130, loss: 0.00747070275247097
step: 140, loss: 0.091057188808918
step: 150, loss: 0.11292432248592377
step: 160, loss: 0.004289062228053808
step: 170, loss: 0.051221564412117004
/usr/local/lib/python3.7/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=0.8307692307692308, f1=0.8706467661691543,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05352722108364105
step: 10, loss: 0.0013716445537284017
step: 20, loss: 0.008519716560840607
step: 30, loss: 0.000658916134852916
step: 40, loss: 0.002000552136451006
step: 50, loss: 0.010083350352942944
step: 60, loss: 0.007297474425286055
step: 70, loss: 0.034770529717206955
step: 80, loss: 0.001875353162176907
step: 90, loss: 0.04638725891709328
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.0007210664916783571
step: 110, loss: 0.028107738122344017
step: 120, loss: 0.010147488676011562
step: 130, loss: 0.004008975811302662
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.004242307972162962
step: 150, loss: 0.0060484749265015125
step: 160, loss: 0.0017838955391198397
step: 170, loss: 0.05814525857567787
/usr/local/lib/python3.7/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=0.8426150121065376, f1=0.8652482269503546,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0127047598361969
step: 10, loss: 0.005102070048451424
step: 20, loss: 0.001470701303333044
step: 30, loss: 0.05356398969888687
step: 40, loss: 0.2840099334716797
step: 50, loss: 0.017903026193380356
step: 60, loss: 0.024330224841833115
step: 70, loss: 0.0012317100772634149
step: 80, loss: 0.00200811680406332
step: 90, loss: 0.09050324559211731
step: 100, loss: 0.03542003408074379
step: 110, loss: 0.058240339159965515
step: 120, loss: 0.00675318343564868
step: 130, loss: 0.0006760115502402186
step: 140, loss: 0.00036239164182916284
step: 150, loss: 0.07413772493600845
step: 160, loss: 0.004766915924847126
step: 170, loss: 0.001246242318302393
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.855072463768116, f1=0.889423076923077,
best_f1=0.889423076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003118712455034256
step: 10, loss: 0.02586132101714611
step: 20, loss: 0.03582154959440231
step: 30, loss: 0.06471970677375793
step: 40, loss: 0.0023538677487522364
step: 50, loss: 0.03847627341747284
step: 60, loss: 0.025568392127752304
step: 70, loss: 0.0013760486617684364
step: 80, loss: 0.038331761956214905
step: 90, loss: 0.0005226263892836869
step: 100, loss: 0.0031447098590433598
step: 110, loss: 0.011809946037828922
step: 120, loss: 0.00359911285340786
step: 130, loss: 0.00236023822799325
step: 140, loss: 0.004157720133662224
step: 150, loss: 0.10392443835735321
step: 160, loss: 0.0009486943599767983
step: 170, loss: 0.0773770660161972

```

```

/usr/local/lib/python3.7/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=0.8456057007125889, f1=0.8695652173913044,
best_f1=0.889423076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03598199412226677
step: 10, loss: 0.003163313725963235
step: 20, loss: 0.003100415226072073
step: 30, loss: 0.0017083886777982116
step: 40, loss: 0.0012532215332612395
step: 50, loss: 0.006100168451666832
step: 60, loss: 0.04421263188123703
step: 70, loss: 0.017291095107793808
step: 80, loss: 0.058587849140167236
step: 90, loss: 0.013011384755373001
step: 100, loss: 0.0014127201866358519
step: 110, loss: 0.01307592075318098
step: 120, loss: 0.003868216648697853
step: 130, loss: 0.015239333733916283
step: 140, loss: 0.013930200599133968
step: 150, loss: 0.009170117788016796
step: 160, loss: 0.011352538131177425
step: 170, loss: 0.02207065373659134
/usr/local/lib/python3.7/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=0.8426395939086295, f1=0.8731707317073171,
best_f1=0.889423076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001437922241166234
step: 10, loss: 0.01319440919905901
step: 20, loss: 0.0008607681374996901
step: 30, loss: 0.11451166868209839
step: 40, loss: 0.0013615402858704329
step: 50, loss: 0.000592939555644989
step: 60, loss: 0.008976300247013569
step: 70, loss: 0.0004596822545863688
step: 80, loss: 0.00101671670563519
step: 90, loss: 0.009076149202883244
step: 100, loss: 0.0009364458383060992
step: 110, loss: 0.021911589428782463
step: 120, loss: 0.021371452137827873
step: 130, loss: 0.036279868334531784

```

```

step: 140, loss: 0.0031656718347221613
step: 150, loss: 0.008512627333402634
step: 160, loss: 0.044147688895463943
step: 170, loss: 0.001724510919302702
/usr/local/lib/python3.7/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=0.8404255319148938, f1=0.8877551020408163,
best_f1=0.889423076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00197223830036819
step: 10, loss: 0.0005847982247360051
step: 20, loss: 0.0026974081993103027
step: 30, loss: 0.05036264285445213
step: 40, loss: 0.001243950449861586
step: 50, loss: 0.020432867109775543
step: 60, loss: 0.0009814569493755698
step: 70, loss: 0.11297915875911713
step: 80, loss: 0.0009771873010322452
step: 90, loss: 0.004193961154669523
step: 100, loss: 0.011812024749815464
step: 110, loss: 0.00046364229638129473
step: 120, loss: 0.023969998583197594
step: 130, loss: 0.00036663946229964495
step: 140, loss: 0.024402089416980743
step: 150, loss: 0.0006598714389838278
step: 160, loss: 0.009149382822215557
step: 170, loss: 0.0003482913307379931
/usr/local/lib/python3.7/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=0.8481675392670156, f1=0.8877805486284289,
best_f1=0.889423076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011866986751556396
step: 10, loss: 0.0008773576118983328
step: 20, loss: 0.001782419509254396
step: 30, loss: 0.008389749564230442
step: 40, loss: 0.003311724402010441
step: 50, loss: 0.00038788188248872757
step: 60, loss: 0.0024536247365176678
step: 70, loss: 0.009792294353246689
step: 80, loss: 0.0003930175444111228
step: 90, loss: 0.00028760204440914094

```

```

step: 100, loss: 0.0013009619433432817
step: 110, loss: 0.021054040640592575
step: 120, loss: 0.001643726951442659
step: 130, loss: 0.1210077702999115
step: 140, loss: 0.03764251247048378
step: 150, loss: 0.03379780799150467
step: 160, loss: 0.0007036264869384468
step: 170, loss: 0.05244353786110878
/usr/local/lib/python3.7/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=0.8514851485148516, f1=0.8782816229116946,
best_f1=0.889423076923077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035809827386401594
step: 10, loss: 0.0006415274692699313
step: 20, loss: 0.000515255203936249
step: 30, loss: 0.0013322425074875355
step: 40, loss: 0.0003719277447089553
step: 50, loss: 0.0003192858712282032
step: 60, loss: 0.0005823928513564169
step: 70, loss: 0.061764393001794815
step: 80, loss: 0.004331403877586126
step: 90, loss: 0.009952993132174015
step: 100, loss: 0.028038958087563515
step: 110, loss: 0.0008500184630975127
step: 120, loss: 0.0002946292224805802
step: 130, loss: 0.0008964926819317043
step: 140, loss: 0.02152494341135025
step: 150, loss: 0.00032445768010802567
step: 160, loss: 0.0004148848238401115
step: 170, loss: 0.0008243889315053821
/usr/local/lib/python3.7/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=0.8519480519480519, f1=0.8793969849246231,
best_f1=0.889423076923077
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 223.64it/s]
load_f1 = 0.4772413793103447
real_f1 = 0.46469833119383824
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 134.24it/s]

```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

```

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

```

```

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

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

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 520kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 2.39MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 1.82MB/s]
Downloading: 100% 501M/501M [00:08<00:00, 57.8MB/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.")
step: 0, loss: 0.5585409998893738
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4479609727859497
step: 20, loss: 0.47656768560409546
step: 30, loss: 0.16349348425865173
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.12789885699748993
step: 50, loss: 0.08718784153461456
step: 60, loss: 0.0627484992146492
step: 70, loss: 0.1625131368637085
step: 80, loss: 0.02822524681687355
step: 90, loss: 0.08259640634059906
step: 100, loss: 0.15459725260734558
step: 110, loss: 0.17773421108722687
step: 120, loss: 0.037919607013463974
step: 130, loss: 0.02520163729786873
step: 140, loss: 0.027222732082009315
step: 150, loss: 0.1438496857881546
step: 160, loss: 0.011608911678195
step: 170, loss: 0.0900685042142868
step: 180, loss: 0.0656827837228775
step: 190, loss: 0.3481449484825134
step: 200, loss: 0.03562541306018829
step: 210, loss: 0.04067953675985336
step: 220, loss: 0.10438183695077896
step: 230, loss: 0.0038431689608842134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9627959413754228, f1=0.9623717217787913,
best_f1=0.9623717217787913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004411256406456232
step: 10, loss: 0.1506527066230774
step: 20, loss: 0.05771924555301666
step: 30, loss: 0.012658203952014446
step: 40, loss: 0.040094196796417236
step: 50, loss: 0.010577218607068062
step: 60, loss: 0.018439944833517075
step: 70, loss: 0.0021214098669588566

```

```

step: 80, loss: 0.013295897282660007
step: 90, loss: 0.017625857144594193
step: 100, loss: 0.015443358570337296
step: 110, loss: 0.037450820207595825
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 120, loss: 0.10275203734636307
step: 130, loss: 0.04892390966415405
step: 140, loss: 0.009983167052268982
step: 150, loss: 0.16328038275241852
step: 160, loss: 0.0525825172662735
step: 170, loss: 0.0027925886679440737
step: 180, loss: 0.004219682887196541
step: 190, loss: 0.016593987122178078
step: 200, loss: 0.02729925699532032
step: 210, loss: 0.03157010301947594
step: 220, loss: 0.006014808546751738
step: 230, loss: 0.010846163146197796
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9722530521642618, f1=0.9654403567447045,
best_f1=0.9654403567447045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003582240315154195
step: 10, loss: 0.01342077273875475
step: 20, loss: 0.00269341841340065
step: 30, loss: 0.004133650101721287
step: 40, loss: 0.05225486680865288
step: 50, loss: 0.031182631850242615
step: 60, loss: 0.012144672684371471
step: 70, loss: 0.008460110053420067
step: 80, loss: 0.11989393085241318
step: 90, loss: 0.01398811861872673
step: 100, loss: 0.003329257946461439
step: 110, loss: 0.06142976135015488
step: 120, loss: 0.0028871544636785984
step: 130, loss: 0.00976966880261898
step: 140, loss: 0.001693323371000588
step: 150, loss: 0.011878821067512035
step: 160, loss: 0.0013827835209667683
step: 170, loss: 0.000985765946097672
step: 180, loss: 0.011342584155499935

```

```

step: 190, loss: 0.023035556077957153
step: 200, loss: 0.014278572984039783
step: 210, loss: 0.0077951038256287575
step: 220, loss: 0.1006995290517807
step: 230, loss: 0.0029380912892520428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753914988814317, f1=0.9641255605381166,
best_f1=0.9641255605381166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012008978053927422
step: 10, loss: 0.0008990142960101366
step: 20, loss: 0.002930688438937068
step: 30, loss: 0.002712762448936701
step: 40, loss: 0.06340226531028748
step: 50, loss: 0.05131198838353157
step: 60, loss: 0.01001882366836071
step: 70, loss: 0.0040076253935694695
step: 80, loss: 0.020073965191841125
step: 90, loss: 0.089272640645504
step: 100, loss: 0.026303503662347794
step: 110, loss: 0.0003820470883511007
step: 120, loss: 0.008012393489480019
step: 130, loss: 0.013326904736459255
step: 140, loss: 0.05245829373598099
step: 150, loss: 0.0021482177544385195
step: 160, loss: 0.0037831219378858805
step: 170, loss: 0.01256315503269434
step: 180, loss: 0.005740619730204344
step: 190, loss: 0.0025310388300567865
step: 200, loss: 0.03250092267990112
step: 210, loss: 0.006964176893234253
step: 220, loss: 0.001390937715768814
step: 230, loss: 0.0007403510389849544
/usr/local/lib/python3.7/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=0.9730941704035874, f1=0.9706546275395034,
best_f1=0.9641255605381166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013672136701643467
step: 10, loss: 0.02613447979092598
step: 20, loss: 0.038354985415935516
step: 30, loss: 0.003283090889453888
step: 40, loss: 0.006660830229520798
step: 50, loss: 0.0030086126644164324
step: 60, loss: 0.013460497371852398
step: 70, loss: 0.002424921840429306
step: 80, loss: 0.07296543568372726
step: 90, loss: 0.05887584760785103
step: 100, loss: 0.00033986152266152203
step: 110, loss: 0.0034400280565023422
step: 120, loss: 0.0017181371804326773
step: 130, loss: 0.00031805472099222243
step: 140, loss: 0.00700764637440443
step: 150, loss: 0.015693500638008118
step: 160, loss: 0.012623663991689682
step: 170, loss: 0.0007929775165393949
step: 180, loss: 0.01386401429772377
step: 190, loss: 0.011135175824165344
step: 200, loss: 0.007049520965665579
step: 210, loss: 0.03770580142736435
step: 220, loss: 0.0027998120058327913
step: 230, loss: 0.03369409218430519
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9799107142857142, f1=0.9763779527559054,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007570014800876379
step: 10, loss: 0.0028100779745727777
step: 20, loss: 0.0005762066575698555
step: 30, loss: 0.0006132987327873707
step: 40, loss: 0.00018105271738022566
step: 50, loss: 0.00513994786888361
step: 60, loss: 0.0056342752650380135
step: 70, loss: 0.18787293136119843
step: 80, loss: 0.0010462728096172214
step: 90, loss: 0.005076193250715733
step: 100, loss: 0.0005193339893594384

```

```

step: 110, loss: 0.022448336705565453
step: 120, loss: 0.0014223720645532012
step: 130, loss: 0.0029690980445593596
step: 140, loss: 0.0006797860260121524
step: 150, loss: 0.002598196268081665
step: 160, loss: 0.0012475713156163692
step: 170, loss: 0.001381515059620142
step: 180, loss: 0.0013660260010510683
step: 190, loss: 0.00048132752999663353
step: 200, loss: 0.0009178614127449691
step: 210, loss: 0.00016847572987899184
step: 220, loss: 0.004426160827279091
step: 230, loss: 0.0003791142371483147
/usr/local/lib/python3.7/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=0.9785310734463276, f1=0.9750566893424036,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007344732875935733
step: 10, loss: 0.0003045230114366859
step: 20, loss: 0.00033931192592717707
step: 30, loss: 0.00010093981836689636
step: 40, loss: 0.00022624807024840266
step: 50, loss: 0.00023164933372754604
step: 60, loss: 0.0002189444930991158
step: 70, loss: 0.0004675121745094657
step: 80, loss: 0.0003009157080668956
step: 90, loss: 0.00021233715233393013
step: 100, loss: 0.00016345834592357278
step: 110, loss: 0.00017023553664330393
step: 120, loss: 0.001230889931321144
step: 130, loss: 0.003993373829871416
step: 140, loss: 7.829837704775855e-05
step: 150, loss: 0.008074553683400154
step: 160, loss: 0.00018716484191827476
step: 170, loss: 0.0002791076258290559
step: 180, loss: 0.00012925911869388074
step: 190, loss: 0.0002695170696824789
step: 200, loss: 0.002806413220241666
step: 210, loss: 0.002236793516203761
step: 220, loss: 0.0012936112470924854
step: 230, loss: 0.0035369531251490116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820224719101124, f1=0.9750566893424036,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004359900485724211
step: 10, loss: 0.0049345423467457294
step: 20, loss: 0.003347745630890131
step: 30, loss: 0.010012284852564335
step: 40, loss: 0.0009227698901668191
step: 50, loss: 0.00035412723082117736
step: 60, loss: 0.0003176293394062668
step: 70, loss: 0.0001162242260761559
step: 80, loss: 0.0013363598845899105
step: 90, loss: 0.002415427239611745
step: 100, loss: 0.000311005802359432
step: 110, loss: 0.0010161908576264977
step: 120, loss: 0.0010611001634970307
step: 130, loss: 0.007979908026754856
step: 140, loss: 0.0003854141978081316
step: 150, loss: 0.253693163394928
step: 160, loss: 0.00027853387291543186
step: 170, loss: 0.013986891135573387
step: 180, loss: 0.00039461307460442185
step: 190, loss: 0.0005274647264741361
step: 200, loss: 0.0018018027767539024
step: 210, loss: 0.01855422370135784
step: 220, loss: 0.0004608457093127072
step: 230, loss: 0.00045684949145652354
/usr/local/lib/python3.7/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=0.9810901001112348, f1=0.9709821428571428,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004806302487850189
step: 10, loss: 0.00024840107653290033
step: 20, loss: 0.0004803628253284842
step: 30, loss: 0.00048474996583536267
step: 40, loss: 0.0012099886080250144
step: 50, loss: 0.0008150585927069187
step: 60, loss: 0.0004297671257518232

```

```

step: 70, loss: 0.050985757261514664
step: 80, loss: 0.00033796942443586886
step: 90, loss: 0.008424879983067513
step: 100, loss: 0.0003863350721076131
step: 110, loss: 0.0003210111754015088
step: 120, loss: 0.028732798993587494
step: 130, loss: 0.0018956454005092382
step: 140, loss: 0.0003481896419543773
step: 150, loss: 0.0004497393383644521
step: 160, loss: 0.0008159317076206207
step: 170, loss: 0.0020284082274883986
step: 180, loss: 0.00049617700278759
step: 190, loss: 0.0001623657444724813
step: 200, loss: 0.006140848156064749
step: 210, loss: 0.0005589410429820418
step: 220, loss: 0.0002461228286847472
step: 230, loss: 8.877224900061265e-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=0.9810901001112348, f1=0.9753363228699552,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019937558681704104
step: 10, loss: 0.00022345077013596892
step: 20, loss: 0.0001533313188701868
step: 30, loss: 0.00013093608140479773
step: 40, loss: 0.0007602378609590232
step: 50, loss: 0.000685730017721653
step: 60, loss: 0.0003090204845648259
step: 70, loss: 0.0014558910625055432
step: 80, loss: 0.0011195589322596788
step: 90, loss: 0.00022846314823254943
step: 100, loss: 0.0002665301435627043
step: 110, loss: 0.0010948176495730877
step: 120, loss: 0.022810062393546104
step: 130, loss: 0.0003786641755141318
step: 140, loss: 0.0001869612606242299
step: 150, loss: 0.004833802580833435
step: 160, loss: 0.0001738392311381176
step: 170, loss: 0.0002083254512399435
step: 180, loss: 0.0014703032793477178
step: 190, loss: 0.0015478839632123709
step: 200, loss: 0.0018090083030983806
step: 210, loss: 0.0001412531128153205
step: 220, loss: 0.0007785193738527596

```

```

step: 230, loss: 0.00010176272189710289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865470852017937, f1=0.9785794813979707,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010008020763052627
step: 10, loss: 0.00035758124431595206
step: 20, loss: 0.00019689268083311617
step: 30, loss: 0.0001447526883566752
step: 40, loss: 3.30555449181702e-05
step: 50, loss: 0.0020088651217520237
step: 60, loss: 0.00016182013496290892
step: 70, loss: 0.00011693496344378218
step: 80, loss: 0.0005417991778813303
step: 90, loss: 0.17939352989196777
step: 100, loss: 0.0010128978174179792
step: 110, loss: 0.003874750342220068
step: 120, loss: 0.00033746642293408513
step: 130, loss: 0.00027006681193597615
step: 140, loss: 0.0003497617435641587
step: 150, loss: 0.00017438281793147326
step: 160, loss: 0.0002416156348772347
step: 170, loss: 0.001986423274502158
step: 180, loss: 0.0016125283436849713
step: 190, loss: 0.00011818340863101184
step: 200, loss: 0.00025498069589957595
step: 210, loss: 0.0007239190745167434
step: 220, loss: 0.0003227928245905787
step: 230, loss: 0.000369417539332062
/usr/local/lib/python3.7/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=0.9843400447427293, f1=0.9764309764309763,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048651121323928237
step: 10, loss: 0.00040340720443055034
step: 20, loss: 0.001043518423102796

```



```

step: 30, loss: 0.03358427435159683
step: 40, loss: 0.0002647133660502732
step: 50, loss: 0.0003536874137353152
step: 60, loss: 0.00019498106848914176
step: 70, loss: 0.00016623202827759087
step: 80, loss: 0.0001097179701901041
step: 90, loss: 0.004302520304918289
step: 100, loss: 0.00015806953888386488
step: 110, loss: 9.565251093590632e-05
step: 120, loss: 0.00022072248975746334
step: 130, loss: 0.00013798322470393032
step: 140, loss: 0.00031184169347397983
step: 150, loss: 0.00024636846501380205
step: 160, loss: 0.00027239619521424174
step: 170, loss: 0.0003797845565713942
step: 180, loss: 0.00020870944717898965
step: 190, loss: 0.002334315562620759
step: 200, loss: 0.00018684992392081767
step: 210, loss: 0.0009263508836738765
step: 220, loss: 0.015325737185776234
step: 230, loss: 0.0002773209707811475
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9785310734463276,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040068512316793203
step: 10, loss: 0.0031325866002589464
step: 20, loss: 0.0002686090883798897
step: 30, loss: 0.00012231399887241423
step: 40, loss: 0.00820744689553976
step: 50, loss: 0.00045598295400850475
step: 60, loss: 0.00017875051707960665
step: 70, loss: 0.00015169060498010367
step: 80, loss: 0.0001886265235953033
step: 90, loss: 0.00013061003119219095
step: 100, loss: 0.0008418122888542712
step: 110, loss: 0.0005435257335193455
step: 120, loss: 0.0003113684360869229
step: 130, loss: 0.00023573910584673285
step: 140, loss: 0.000573948142118752
step: 150, loss: 4.65833509224467e-05
step: 160, loss: 0.005499429535120726
step: 170, loss: 0.000349146721418947
step: 180, loss: 0.03462295979261398

```

```

step: 190, loss: 0.00021338442456908524
step: 200, loss: 7.441000343533233e-05
step: 210, loss: 0.0010786026250571012
step: 220, loss: 0.00014953850768506527
step: 230, loss: 0.00023951599723659456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876265466816648, f1=0.983050847457627,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011430263111833483
step: 10, loss: 0.00010706399189075455
step: 20, loss: 0.00020018426585011184
step: 30, loss: 0.00023634533863514662
step: 40, loss: 0.0004001825291197747
step: 50, loss: 0.00020967585442122072
step: 60, loss: 0.00022105930838733912
step: 70, loss: 0.0001723864406812936
step: 80, loss: 0.0005684954230673611
step: 90, loss: 0.00034186517586931586
step: 100, loss: 0.0008838950889185071
step: 110, loss: 0.00021265377290546894
step: 120, loss: 7.324146281462163e-05
step: 130, loss: 0.0013132102321833372
step: 140, loss: 0.0001380109169986099
step: 150, loss: 0.0002712627174332738
step: 160, loss: 0.0001601896365173161
step: 170, loss: 0.0010273198131471872
step: 180, loss: 0.0001387779921060428
step: 190, loss: 0.0001077064371202141
step: 200, loss: 0.0006498225266113877
step: 210, loss: 0.00013825474889017642
step: 220, loss: 0.00014094215293880552
step: 230, loss: 8.553636871511117e-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=0.9853768278965129, f1=0.9819004524886877,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013541012071073055
step: 10, loss: 0.00011475652718218043
step: 20, loss: 0.000135682916152291
step: 30, loss: 0.00031340529676526785
step: 40, loss: 6.461209704866633e-05
step: 50, loss: 0.00010024240327766165
step: 60, loss: 0.014501716941595078
step: 70, loss: 0.00013411165855359286
step: 80, loss: 0.00011108470062026754
step: 90, loss: 8.796157635515556e-05
step: 100, loss: 6.400574784493074e-05
step: 110, loss: 0.00011320797784719616
step: 120, loss: 0.040704671293497086
step: 130, loss: 7.118190114852041e-05
step: 140, loss: 0.0001390371035085991
step: 150, loss: 0.0003426697221584618
step: 160, loss: 0.0001353968254989013
step: 170, loss: 5.5468066420871764e-05
step: 180, loss: 0.0001150181851699017
step: 190, loss: 0.01828501932322979
step: 200, loss: 0.00010777825809782371
step: 210, loss: 0.00011634820839390159
step: 220, loss: 0.004999329801648855
step: 230, loss: 0.00013322725135367364
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.984304932735426, f1=0.9807909604519773,
best_f1=0.983050847457627
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:14, 171.76it/s]
load_f1 = 0.9865168539325843
real_f1 = 0.9853768278965129
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 [00:27, 161.42it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.6220571398735046

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

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

step: 10, loss: 0.4626280963420868

step: 20, loss: 0.26652851700782776

step: 30, loss: 0.3468167185783386

step: 40, loss: 0.2596467435359955

step: 50, loss: 0.20150010287761688

step: 60, loss: 0.11904998123645782

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

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

step: 70, loss: 0.34612521529197693

step: 80, loss: 0.1582440882921219

step: 90, loss: 0.18110418319702148

```
step: 100, loss: 0.20499184727668762
step: 110, loss: 0.14274635910987854
step: 120, loss: 0.19262324273586273
step: 130, loss: 0.16089962422847748
step: 140, loss: 0.1366586834192276
step: 150, loss: 0.017306264489889145
step: 160, loss: 0.11653464287519455
step: 170, loss: 0.10048340260982513
step: 180, loss: 0.03785916045308113
step: 190, loss: 0.048510368913412094
step: 200, loss: 0.07477514445781708
step: 210, loss: 0.02336486056447029
step: 220, loss: 0.14779475331306458
step: 230, loss: 0.1710854172706604
step: 240, loss: 0.08538270741701126
step: 250, loss: 0.03117193654179573
step: 260, loss: 0.09383100271224976
step: 270, loss: 0.3879292607307434
step: 280, loss: 0.04976237937808037
step: 290, loss: 0.03950022533535957
step: 300, loss: 0.10343682765960693
step: 310, loss: 0.06501466780900955
step: 320, loss: 0.08078210800886154
step: 330, loss: 0.14627037942409515
step: 340, loss: 0.2792011499404907
step: 350, loss: 0.06782636791467667
step: 360, loss: 0.08378702402114868
step: 370, loss: 0.076566182076931
step: 380, loss: 0.2617242634296417
step: 390, loss: 0.036632537841796875
step: 400, loss: 0.04871857166290283
step: 410, loss: 0.33778640627861023
step: 420, loss: 0.01085746381431818
step: 430, loss: 0.032009370625019073
step: 440, loss: 0.07187193632125854
step: 450, loss: 0.043000727891922
step: 460, loss: 0.041885558515787125
step: 470, loss: 0.08298017829656601
step: 480, loss: 0.1800771802663803
step: 490, loss: 0.058772940188646317
step: 500, loss: 0.09405848383903503
step: 510, loss: 0.054598692804574966
step: 520, loss: 0.01721426099538803
step: 530, loss: 0.019758479669690132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.937962962962963, f1=0.9339491916859123,
best_f1=0.9339491916859123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

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

```
step: 0, loss: 0.022134048864245415
step: 10, loss: 0.15479543805122375
step: 20, loss: 0.09500069171190262
step: 30, loss: 0.020609308034181595
step: 40, loss: 0.06937026977539062
step: 50, loss: 0.026062849909067154
step: 60, loss: 0.021925345063209534
step: 70, loss: 0.06280285865068436
step: 80, loss: 0.03293908014893532
step: 90, loss: 0.004362322855740786
step: 100, loss: 0.11475042998790741
step: 110, loss: 0.006541657727211714
step: 120, loss: 0.0511583611369133
step: 130, loss: 0.007238083053380251
step: 140, loss: 0.01734072156250477
step: 150, loss: 0.004719840362668037
step: 160, loss: 0.023563163354992867
step: 170, loss: 0.059913016855716705
step: 180, loss: 0.029545947909355164
step: 190, loss: 0.028042474761605263
step: 200, loss: 0.12082647532224655
step: 210, loss: 0.023798424750566483
step: 220, loss: 0.0015553999692201614
step: 230, loss: 0.059583425521850586
step: 240, loss: 0.02879883535206318
step: 250, loss: 0.03348800539970398
step: 260, loss: 0.13612490892410278
step: 270, loss: 0.010139341466128826
step: 280, loss: 0.03621207922697067
step: 290, loss: 0.05053841322660446
step: 300, loss: 0.052870530635118484
step: 310, loss: 0.023411480709910393
step: 320, loss: 0.029206939041614532
step: 330, loss: 0.04060026630759239
step: 340, loss: 0.12559543550014496
step: 350, loss: 0.005786160472780466
step: 360, loss: 0.06359905004501343
step: 370, loss: 0.010941724292933941
step: 380, loss: 0.08406448364257812
step: 390, loss: 0.003005056641995907
```

```

step: 400, loss: 0.10672584921121597
step: 410, loss: 0.04738518223166466
step: 420, loss: 0.016387278214097023
step: 430, loss: 0.18570218980312347
step: 440, loss: 0.00525784632191062
step: 450, loss: 0.06615812331438065
step: 460, loss: 0.09186585992574692
step: 470, loss: 0.06490188091993332
step: 480, loss: 0.003471317468211055
step: 490, loss: 0.03117303177714348
step: 500, loss: 0.011477747932076454
step: 510, loss: 0.03665812686085701
step: 520, loss: 0.25511205196380615
step: 530, loss: 0.10194824635982513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.954060324825986, f1=0.9547553093259464,
best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17956344783306122
step: 10, loss: 0.09091513603925705
step: 20, loss: 0.003375398926436901
step: 30, loss: 0.0823824405670166
step: 40, loss: 0.06862549483776093
step: 50, loss: 0.009791580028831959
step: 60, loss: 0.03050249069929123
step: 70, loss: 0.008589518256485462
step: 80, loss: 0.019843509420752525
step: 90, loss: 0.004342060070484877
step: 100, loss: 0.016434499993920326
step: 110, loss: 0.00935788918286562
step: 120, loss: 0.18550170958042145
step: 130, loss: 0.04054898023605347
step: 140, loss: 0.06226503476500511
step: 150, loss: 0.012812057510018349
step: 160, loss: 0.0042368280701339245
step: 170, loss: 0.0582955926656723
step: 180, loss: 0.012628485448658466
step: 190, loss: 0.013984227553009987
step: 200, loss: 0.018035244196653366
step: 210, loss: 0.025990167632699013

```



```

step: 220, loss: 0.16977155208587646
step: 230, loss: 0.13693563640117645
step: 240, loss: 0.11155788600444794
step: 250, loss: 0.08258332312107086
step: 260, loss: 0.146664097905159
step: 270, loss: 0.02808353118598461
step: 280, loss: 0.06337558478116989
step: 290, loss: 0.04025822505354881
step: 300, loss: 0.19724339246749878
step: 310, loss: 0.0945829376578331
step: 320, loss: 0.008044252172112465
step: 330, loss: 0.010585475713014603
step: 340, loss: 0.015756910666823387
step: 350, loss: 0.05784076452255249
step: 360, loss: 0.012008203193545341
step: 370, loss: 0.04536915197968483
step: 380, loss: 0.006822832860052586
step: 390, loss: 0.003506760112941265
step: 400, loss: 0.11072894185781479
step: 410, loss: 0.024643706157803535
step: 420, loss: 0.009783321060240269
step: 430, loss: 0.048137810081243515
step: 440, loss: 0.15968698263168335
step: 450, loss: 0.009622008539736271
step: 460, loss: 0.0807073786854744
step: 470, loss: 0.034710489213466644
step: 480, loss: 0.053507495671510696
step: 490, loss: 0.04398760944604874
step: 500, loss: 0.023576153442263603
step: 510, loss: 0.026297202333807945
step: 520, loss: 0.06987935304641724
step: 530, loss: 0.022321229800581932
/usr/local/lib/python3.7/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=0.9461966604823748, f1=0.95, best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022596152499318123
step: 10, loss: 0.0028470843099057674
step: 20, loss: 0.004982605576515198
step: 30, loss: 0.1819055825471878
step: 40, loss: 0.009557968936860561
step: 50, loss: 0.04146171361207962
step: 60, loss: 0.003965109121054411
step: 70, loss: 0.044732943177223206
step: 80, loss: 0.08420481532812119

```

```
step: 90, loss: 0.0532212071120739
step: 100, loss: 0.006689711939543486
step: 110, loss: 0.02845119498670101
step: 120, loss: 0.003971385303884745
step: 130, loss: 0.05038256198167801
step: 140, loss: 0.09192594885826111
step: 150, loss: 0.016527103260159492
step: 160, loss: 0.03899814933538437
step: 170, loss: 0.04731959477066994
step: 180, loss: 0.11199651658535004
step: 190, loss: 0.026843639090657234
step: 200, loss: 0.010044220834970474
step: 210, loss: 0.0013900598278269172
step: 220, loss: 0.030657587572932243
step: 230, loss: 0.003312179585918784
step: 240, loss: 0.08573035895824432
step: 250, loss: 0.0676264837384224
step: 260, loss: 0.017920570448040962
step: 270, loss: 0.10164779424667358
step: 280, loss: 0.02035265974700451
step: 290, loss: 0.004839980974793434
step: 300, loss: 0.001939311041496694
step: 310, loss: 0.002564616734161973
step: 320, loss: 0.05529644712805748
step: 330, loss: 0.01869182474911213
step: 340, loss: 0.005219633225351572
step: 350, loss: 0.16672073304653168
step: 360, loss: 0.09591741859912872
step: 370, loss: 0.004411250352859497
step: 380, loss: 0.03694285452365875
step: 390, loss: 0.0012686483096331358
step: 400, loss: 0.004529115278273821
step: 410, loss: 0.0018910749349743128
step: 420, loss: 0.07530365884304047
step: 430, loss: 0.02922820672392845
step: 440, loss: 0.0030367542058229446
step: 450, loss: 0.009548122994601727
step: 460, loss: 0.02021385356783867
step: 470, loss: 0.0027961714658886194
step: 480, loss: 0.0067482562735676765
step: 490, loss: 0.0320211723446846
step: 500, loss: 0.005261437967419624
step: 510, loss: 0.020170124247670174
step: 520, loss: 0.05375562608242035
step: 530, loss: 0.08134731650352478
/usr/local/lib/python3.7/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=0.9510032664489033, f1=0.9471733086190918,  
best_f1=0.9547553093259464  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0034798127599060535  
step: 10, loss: 0.011980244889855385  
step: 20, loss: 0.005382912699133158  
step: 30, loss: 0.010982303880155087  
step: 40, loss: 0.0010129421716555953  
step: 50, loss: 0.17500527203083038  
step: 60, loss: 0.025476107373833656  
step: 70, loss: 0.0020603490993380547  
step: 80, loss: 0.005925409961491823  
step: 90, loss: 0.12165307253599167  
step: 100, loss: 0.056351255625486374  
step: 110, loss: 0.0028212612960487604  
step: 120, loss: 0.14273281395435333  
step: 130, loss: 0.007828240282833576  
step: 140, loss: 0.04605681821703911  
step: 150, loss: 0.06982804834842682  
step: 160, loss: 0.05199345201253891  
step: 170, loss: 0.10812518000602722  
step: 180, loss: 0.04167800024151802  
step: 190, loss: 0.028298290446400642  
step: 200, loss: 0.019513361155986786  
step: 210, loss: 0.005004600156098604  
step: 220, loss: 0.002010623924434185  
step: 230, loss: 0.0038463156670331955  
step: 240, loss: 0.00195152941159904  
step: 250, loss: 0.08381538093090057  
step: 260, loss: 0.004028339870274067  
step: 270, loss: 0.03849392384290695  
step: 280, loss: 0.0026131970807909966  
step: 290, loss: 0.0450429692864418  
step: 300, loss: 0.0209529846906662  
step: 310, loss: 0.04975868761539459  
step: 320, loss: 0.007493270561099052  
step: 330, loss: 0.08817914873361588  
step: 340, loss: 0.011285820044577122  
step: 350, loss: 0.0032367261592298746  
step: 360, loss: 0.000617857847828418  
step: 370, loss: 0.0004746159538626671  
step: 380, loss: 0.0014326982200145721  
step: 390, loss: 0.001196464290842414  
step: 400, loss: 0.01938738115131855  
step: 410, loss: 0.13612079620361328  
step: 420, loss: 0.1645815372467041
```

```
step: 430, loss: 0.0027388387825340033
step: 440, loss: 0.0016234948998317122
step: 450, loss: 0.027462782338261604
step: 460, loss: 0.19566111266613007
step: 470, loss: 0.011519302614033222
step: 480, loss: 0.028685076162219048
step: 490, loss: 0.011737039312720299
step: 500, loss: 0.0064319586381316185
step: 510, loss: 0.004835622385144234
step: 520, loss: 0.16866004467010498
step: 530, loss: 0.06245416775345802
/usr/local/lib/python3.7/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=0.9536178107606679, f1=0.9479981592268752,
best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03516734391450882
step: 10, loss: 0.00422839168459177
step: 20, loss: 0.019729051738977432
step: 30, loss: 0.0012632700381800532
step: 40, loss: 0.042167939245700836
step: 50, loss: 0.0015143216587603092
step: 60, loss: 0.0396009162068367
step: 70, loss: 0.0009823995642364025
step: 80, loss: 0.0012967631919309497
step: 90, loss: 0.043218113481998444
step: 100, loss: 0.0013564530527219176
step: 110, loss: 0.004257440101355314
step: 120, loss: 0.025953633710741997
step: 130, loss: 0.010816952213644981
step: 140, loss: 0.0015460068825632334
step: 150, loss: 0.00220690225251019
step: 160, loss: 0.012583786621689796
step: 170, loss: 0.0019876817241311073
step: 180, loss: 0.07078836113214493
step: 190, loss: 0.03999420627951622
step: 200, loss: 0.002288335934281349
step: 210, loss: 0.002602630527690053
step: 220, loss: 0.02637449838221073
step: 230, loss: 0.004642650950700045
step: 240, loss: 0.04653489962220192
step: 250, loss: 0.026980943977832794
step: 260, loss: 0.0221877284348011
step: 270, loss: 0.004451894201338291
step: 280, loss: 0.0022932186257094145
```

```

step: 290, loss: 0.05008920282125473
step: 300, loss: 0.0024966192431747913
step: 310, loss: 0.11174637079238892
step: 320, loss: 0.0008631569216959178
step: 330, loss: 0.0014753342838957906
step: 340, loss: 0.001012918190099299
step: 350, loss: 0.014894981868565083
step: 360, loss: 0.006771457847207785
step: 370, loss: 0.011699521914124489
step: 380, loss: 0.0005355894099920988
step: 390, loss: 0.008412335999310017
step: 400, loss: 0.00040215003537014127
step: 410, loss: 0.000543381494935602
step: 420, loss: 0.019122034311294556
step: 430, loss: 0.002988958265632391
step: 440, loss: 0.0011750375851988792
step: 450, loss: 0.11954040825366974
step: 460, loss: 0.002261321060359478
step: 470, loss: 0.0031018482986837626
step: 480, loss: 0.014264988712966442
step: 490, loss: 0.013276626355946064
step: 500, loss: 0.003492870833724737
step: 510, loss: 0.12209953367710114
step: 520, loss: 0.0064332690089941025
step: 530, loss: 0.0017771226121112704
/usr/local/lib/python3.7/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=0.9493610979649787, f1=0.9456928838951311,
best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001906294608488679
step: 10, loss: 0.001226037391461432
step: 20, loss: 0.004945417866110802
step: 30, loss: 0.006244598422199488
step: 40, loss: 0.009027059189975262
step: 50, loss: 0.0074426899664103985
step: 60, loss: 0.041007377207279205
step: 70, loss: 0.024951757863163948
step: 80, loss: 0.002049219561740756
step: 90, loss: 0.006948438007384539
step: 100, loss: 0.0024326166603714228
step: 110, loss: 0.0008087399764917791
step: 120, loss: 0.0005095581873320043
step: 130, loss: 0.00035754419513978064
step: 140, loss: 0.0012159424368292093

```

```

step: 150, loss: 0.0007652039057575166
step: 160, loss: 0.0016473665600642562
step: 170, loss: 0.005938410758972168
step: 180, loss: 0.10929665714502335
step: 190, loss: 0.017832593992352486
step: 200, loss: 0.001366956508718431
step: 210, loss: 0.0074947490356862545
step: 220, loss: 0.0007162687252275646
step: 230, loss: 0.0005807594861835241
step: 240, loss: 0.0013653385685756803
step: 250, loss: 0.0013915820745751262
step: 260, loss: 0.0016216732328757644
step: 270, loss: 0.0012982861371710896
step: 280, loss: 0.0027467645704746246
step: 290, loss: 0.0026802755892276764
step: 300, loss: 0.000506006006617099
step: 310, loss: 0.0011124603915959597
step: 320, loss: 0.01208737026900053
step: 330, loss: 0.0002428942098049447
step: 340, loss: 0.0011774091981351376
step: 350, loss: 0.0013082429068163037
step: 360, loss: 0.04738859459757805
step: 370, loss: 0.0033760936930775642
step: 380, loss: 0.009041862562298775
step: 390, loss: 0.0072297994047403336
step: 400, loss: 0.015507875010371208
step: 410, loss: 0.0016729366034269333
step: 420, loss: 0.007174242287874222
step: 430, loss: 0.005366507451981306
step: 440, loss: 0.000958563934545964
step: 450, loss: 0.0021475614048540592
step: 460, loss: 0.0007892265566624701
step: 470, loss: 0.12394203245639801
step: 480, loss: 0.002446886384859681
step: 490, loss: 0.002236337633803487
step: 500, loss: 0.0016832235269248486
step: 510, loss: 0.0006403324659913778
step: 520, loss: 0.0014836677582934499
step: 530, loss: 0.0014191963709890842
/usr/local/lib/python3.7/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=0.9463148316651503, f1=0.9481346678798908,
best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1565418392419815

```

step: 10, loss: 0.059413451701402664
step: 20, loss: 0.013411137275397778
step: 30, loss: 0.003965911455452442
step: 40, loss: 0.0002426074497634545
step: 50, loss: 0.0014605928445234895
step: 60, loss: 0.0008020767709240317
step: 70, loss: 0.0017125883605331182
step: 80, loss: 0.0017815626924857497
step: 90, loss: 0.001317458925768733
step: 100, loss: 0.002315725199878216
step: 110, loss: 0.0006078488077037036
step: 120, loss: 0.005550189409404993
step: 130, loss: 0.0008765104576013982
step: 140, loss: 0.0005289882537908852
step: 150, loss: 0.0010677992831915617
step: 160, loss: 0.1209186539053917
step: 170, loss: 0.06054610759019852
step: 180, loss: 0.002580101601779461
step: 190, loss: 0.04683450609445572
step: 200, loss: 0.006283028982579708
step: 210, loss: 0.10376600921154022
step: 220, loss: 0.003586444305256009
step: 230, loss: 0.043931107968091965
step: 240, loss: 0.10287030786275864
step: 250, loss: 0.004077350255101919
step: 260, loss: 0.0010085708927363157
step: 270, loss: 0.015471384860575199
step: 280, loss: 0.00028579786885529757
step: 290, loss: 0.01366041973233223
step: 300, loss: 0.00043416942935436964
step: 310, loss: 0.003104287665337324
step: 320, loss: 0.0026238851714879274
step: 330, loss: 0.0009938105940818787
step: 340, loss: 0.12253013998270035
step: 350, loss: 0.0006505337660200894
step: 360, loss: 0.022020040079951286
step: 370, loss: 0.0033797358628362417
step: 380, loss: 0.0015712715685367584
step: 390, loss: 0.08217630535364151
step: 400, loss: 0.00288764457218349
step: 410, loss: 0.042402077466249466
step: 420, loss: 0.001964577939361334
step: 430, loss: 0.00398401590064168
step: 440, loss: 0.04412742331624031
step: 450, loss: 0.002013010438531637
step: 460, loss: 0.005884903483092785
step: 470, loss: 0.0866934061050415
step: 480, loss: 0.006198082119226456

```

step: 490, loss: 0.056316327303647995
step: 500, loss: 0.004961267579346895
step: 510, loss: 0.03961942344903946
step: 520, loss: 0.0009605613886378706
step: 530, loss: 0.03632698953151703
/usr/local/lib/python3.7/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=0.9483960948396094, f1=0.9443166129774505,
best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012928701471537352
step: 10, loss: 0.016651207581162453
step: 20, loss: 0.0004356523568276316
step: 30, loss: 0.08264881372451782
step: 40, loss: 0.0005628377548418939
step: 50, loss: 0.009179978631436825
step: 60, loss: 0.00420387415215373
step: 70, loss: 0.0007739194552414119
step: 80, loss: 0.046750858426094055
step: 90, loss: 0.07495266199111938
step: 100, loss: 0.0020660972222685814
step: 110, loss: 0.018238786607980728
step: 120, loss: 0.0009232641314156353
step: 130, loss: 0.0005017434596084058
step: 140, loss: 0.000580216059461236
step: 150, loss: 0.0034552672877907753
step: 160, loss: 0.010476961731910706
step: 170, loss: 0.005780437029898167
step: 180, loss: 0.0008135979878716171
step: 190, loss: 0.00029445980908349156
step: 200, loss: 0.00032859580824151635
step: 210, loss: 0.0018935713451355696
step: 220, loss: 0.022949770092964172
step: 230, loss: 0.00022625195560976863
step: 240, loss: 0.000181771902134642
step: 250, loss: 0.000539391883648932
step: 260, loss: 0.0024860436096787453
step: 270, loss: 0.0026262367609888315
step: 280, loss: 0.008486185222864151
step: 290, loss: 0.0029886802658438683
step: 300, loss: 0.0002775181201286614
step: 310, loss: 0.33080366253852844
step: 320, loss: 0.00036292997538112104
step: 330, loss: 0.0011350803542882204
step: 340, loss: 0.009422248229384422

```



```
step: 350, loss: 0.013976921327412128
step: 360, loss: 0.0005667972727678716
step: 370, loss: 0.0004918525810353458
step: 380, loss: 0.0035950487945228815
step: 390, loss: 8.838673966238275e-05
step: 400, loss: 0.08445454388856888
step: 410, loss: 0.15739363431930542
step: 420, loss: 0.00120343582239002
step: 430, loss: 0.004156353883445263
step: 440, loss: 0.008120032027363777
step: 450, loss: 0.08441192656755447
step: 460, loss: 0.003559407312422991
step: 470, loss: 0.0012232217704877257
step: 480, loss: 0.003307420527562499
step: 490, loss: 0.006463410798460245
step: 500, loss: 0.03752698749303818
step: 510, loss: 0.00439953338354826
step: 520, loss: 0.005391254555433989
step: 530, loss: 0.03669652342796326
/usr/local/lib/python3.7/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=0.9496268656716418, f1=0.9432558139534885,
best_f1=0.9547553093259464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005709010292775929
step: 10, loss: 0.0010689503978937864
step: 20, loss: 0.0005848311702720821
step: 30, loss: 0.0070200711488723755
step: 40, loss: 0.005448784679174423
step: 50, loss: 0.003107762197032571
step: 60, loss: 0.0050271255895495415
step: 70, loss: 0.001714398036710918
step: 80, loss: 0.0008119508856907487
step: 90, loss: 0.0041663614101707935
step: 100, loss: 0.0017075338400900364
step: 110, loss: 0.02083691395819187
step: 120, loss: 0.0003212395531591028
step: 130, loss: 0.00011241757601965219
step: 140, loss: 0.03181580826640129
step: 150, loss: 0.0007825085194781423
step: 160, loss: 0.02439050003886223
step: 170, loss: 0.00035619703703559935
step: 180, loss: 0.0016792549286037683
step: 190, loss: 0.002142292447388172
step: 200, loss: 0.0004963835817761719
```

```

step: 210, loss: 0.0017057204386219382
step: 220, loss: 0.0036179067101329565
step: 230, loss: 0.00040206831181421876
step: 240, loss: 0.0015036858385428786
step: 250, loss: 0.00969681702554226
step: 260, loss: 0.005439200438559055
step: 270, loss: 0.024820297956466675
step: 280, loss: 0.003966299816966057
step: 290, loss: 0.002805423690006137
step: 300, loss: 0.014111703261733055
step: 310, loss: 0.041178688406944275
step: 320, loss: 0.05687764286994934
step: 330, loss: 0.012340332381427288
step: 340, loss: 0.004189128056168556
step: 350, loss: 0.008075069636106491
step: 360, loss: 4.865694063482806e-05
step: 370, loss: 0.0003922129690181464
step: 380, loss: 0.01879173144698143
step: 390, loss: 4.8841317038750276e-05
step: 400, loss: 0.0021038162522017956
step: 410, loss: 0.006915663368999958
step: 420, loss: 0.014834277331829071
step: 430, loss: 0.0009260872611775994
step: 440, loss: 0.00030742710805498064
step: 450, loss: 0.0007610578904859722
step: 460, loss: 0.004952242597937584
step: 470, loss: 0.004858986474573612
step: 480, loss: 0.0021205253433436155
step: 490, loss: 0.021071085706353188
step: 500, loss: 0.0016197210643440485
step: 510, loss: 0.00021093557006679475
step: 520, loss: 0.0220600888133049
step: 530, loss: 0.005117427092045546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9558344955834496, f1=0.9458218549127642,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016439781757071614
step: 10, loss: 0.0003914895060006529
step: 20, loss: 0.0004794469859916717

```

step: 30, loss: 0.0028746335301548243
step: 40, loss: 0.00036443761200644076
step: 50, loss: 0.07610417902469635
step: 60, loss: 0.0011132435174658895
step: 70, loss: 0.004063639789819717
step: 80, loss: 0.003784486558288336
step: 90, loss: 0.0012163665378466249
step: 100, loss: 0.006594839505851269
step: 110, loss: 0.0002408490254310891
step: 120, loss: 0.00021243453375063837
step: 130, loss: 0.004180889576673508
step: 140, loss: 0.00014966129674576223
step: 150, loss: 0.0002731101121753454
step: 160, loss: 0.05095764994621277
step: 170, loss: 0.007597845979034901
step: 180, loss: 0.0002220384922111407
step: 190, loss: 0.0002188889484386891
step: 200, loss: 0.005151016637682915
step: 210, loss: 0.002614547498524189
step: 220, loss: 0.13640643656253815
step: 230, loss: 0.0015495165716856718
step: 240, loss: 0.0010790093801915646
step: 250, loss: 0.004688588436692953
step: 260, loss: 0.03609711304306984
step: 270, loss: 0.0019691495690494776
step: 280, loss: 0.0018233164446428418
step: 290, loss: 0.0036904867738485336
step: 300, loss: 0.0002314830489922315
step: 310, loss: 0.024980325251817703
step: 320, loss: 0.003821001620963216
step: 330, loss: 3.786236993619241e-05
step: 340, loss: 0.0074603664688766
step: 350, loss: 0.00028261254192329943
step: 360, loss: 0.0010572158498689532
step: 370, loss: 0.0021233835723251104
step: 380, loss: 0.0013037690659984946
step: 390, loss: 0.0015027327463030815
step: 400, loss: 0.00010335664410376921
step: 410, loss: 0.0029785025399178267
step: 420, loss: 0.0012662550434470177
step: 430, loss: 0.006708802189677954
step: 440, loss: 0.001018901588395238
step: 450, loss: 0.001540318364277482
step: 460, loss: 0.0003015294496435672
step: 470, loss: 0.0012930243974551558
step: 480, loss: 0.002696006093174219
step: 490, loss: 0.0007854228024370968
step: 500, loss: 0.0018123614136129618

```

step: 510, loss: 0.0005118005210533738
step: 520, loss: 0.000841824512463063
step: 530, loss: 0.06837733089923859
/usr/local/lib/python3.7/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=0.9508348794063081, f1=0.9448529411764707,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000624591950327158
step: 10, loss: 0.008647632785141468
step: 20, loss: 0.001586451893672347
step: 30, loss: 0.015103642828762531
step: 40, loss: 0.05012885108590126
step: 50, loss: 0.00021987977379467338
step: 60, loss: 0.0004964562249369919
step: 70, loss: 0.01104719191789627
step: 80, loss: 0.0008629172225482762
step: 90, loss: 0.03172438591718674
step: 100, loss: 0.02026944048702717
step: 110, loss: 0.0001529379078419879
step: 120, loss: 0.0005139141576364636
step: 130, loss: 0.0008794312598183751
step: 140, loss: 6.829834455857053e-05
step: 150, loss: 0.000943389255553484
step: 160, loss: 0.0012610235717147589
step: 170, loss: 0.0004164091369602829
step: 180, loss: 0.006197658833116293
step: 190, loss: 0.0018614373402670026
step: 200, loss: 0.0003528919187374413
step: 210, loss: 9.145514013653155e-06
step: 220, loss: 0.0011119995033368468
step: 230, loss: 5.896097354707308e-05
step: 240, loss: 0.00044789613457396626
step: 250, loss: 0.004048128146678209
step: 260, loss: 0.00013035252050030977
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 270, loss: 0.002008466050028801
step: 280, loss: 0.0010777334682643414
step: 290, loss: 0.001029622508212924
step: 300, loss: 0.0011699422029778361
step: 310, loss: 0.0006424372550100088
step: 320, loss: 0.0014755059964954853
step: 330, loss: 0.0014346682000905275
step: 340, loss: 0.0009914010297507048
step: 350, loss: 3.6565939808497205e-05

```

```

step: 360, loss: 6.173520523589104e-05
step: 370, loss: 8.199031435651705e-05
step: 380, loss: 0.0007370047969743609
step: 390, loss: 0.0024229642003774643
step: 400, loss: 3.515643402351998e-05
step: 410, loss: 6.203104567248374e-05
step: 420, loss: 0.00014095334336161613
step: 430, loss: 8.322314533870667e-05
step: 440, loss: 0.003780281636863947
step: 450, loss: 0.002567332237958908
step: 460, loss: 0.00011377298505976796
step: 470, loss: 0.002898464910686016
step: 480, loss: 0.0003913140099029988
step: 490, loss: 0.01640947535634041
step: 500, loss: 0.004186353180557489
step: 510, loss: 0.020837558433413506
step: 520, loss: 0.0011257166042923927
step: 530, loss: 0.11627226322889328
/usr/local/lib/python3.7/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=0.9512419503219871, f1=0.9495225102319236,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006505689234472811
step: 10, loss: 0.0006977349985390902
step: 20, loss: 0.0003907065838575363
step: 30, loss: 0.00011696215369738638
step: 40, loss: 0.0010318155400454998
step: 50, loss: 0.005226987414062023
step: 60, loss: 0.00034906930522993207
step: 70, loss: 0.011901194229722023
step: 80, loss: 0.0004999330849386752
step: 90, loss: 0.00046235238551162183
step: 100, loss: 9.089165541809052e-05
step: 110, loss: 0.0015975510468706489
step: 120, loss: 0.07574913650751114
step: 130, loss: 0.00038687250344082713
step: 140, loss: 0.0007179511594586074
step: 150, loss: 0.0015868607442826033
step: 160, loss: 0.00033184251515194774
step: 170, loss: 0.007836990989744663
step: 180, loss: 0.002902337582781911
step: 190, loss: 0.001974688144400716
step: 200, loss: 0.00033332881866954267
step: 210, loss: 2.4249078705906868e-05

```

```

step: 220, loss: 0.0013399762101471424
step: 230, loss: 0.0025449609383940697
step: 240, loss: 0.0009452661033719778
step: 250, loss: 0.0009683871758170426
step: 260, loss: 0.00020418927306309342
step: 270, loss: 0.010058343410491943
step: 280, loss: 0.0010034908773377538
step: 290, loss: 0.002223049756139517
step: 300, loss: 0.00018128483498003334
step: 310, loss: 0.00035597910755313933
step: 320, loss: 0.0007305790786631405
step: 330, loss: 0.00018105968774762005
step: 340, loss: 0.0014498198870569468
step: 350, loss: 0.0004764416953548789
step: 360, loss: 0.09226123988628387
step: 370, loss: 0.0005311286076903343
step: 380, loss: 0.0004367568762972951
step: 390, loss: 0.0004754708206746727
step: 400, loss: 0.00022644612181466073
step: 410, loss: 0.00017450182349421084
step: 420, loss: 0.0013809505617246032
step: 430, loss: 0.00017638357530813664
step: 440, loss: 6.618099723709747e-05
step: 450, loss: 0.001152503420598805
step: 460, loss: 0.0012249370338395238
step: 470, loss: 0.003598256967961788
step: 480, loss: 1.4449809896177612e-05
step: 490, loss: 7.575419294880703e-05
step: 500, loss: 0.0015460848808288574
step: 510, loss: 9.781679545994848e-05
step: 520, loss: 0.00010843863128684461
step: 530, loss: 2.1378527890192345e-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=0.9536607970342911, f1=0.9471766848816029,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001652817358262837
step: 10, loss: 0.005978011060506105
step: 20, loss: 0.00011433332110755146
step: 30, loss: 0.010863921605050564
step: 40, loss: 0.0004823802155442536
step: 50, loss: 0.002841286826878786
step: 60, loss: 0.0017845446709543467
step: 70, loss: 0.001105502131395042

```

```
step: 80, loss: 0.00017298395687248558
step: 90, loss: 4.5912081986898556e-05
step: 100, loss: 0.00016160329687409103
step: 110, loss: 3.734184429049492e-05
step: 120, loss: 1.2363562746031675e-05
step: 130, loss: 4.328958675614558e-05
step: 140, loss: 0.0019504689844325185
step: 150, loss: 0.0008165334002114832
step: 160, loss: 0.0001196166267618537
step: 170, loss: 0.014092711731791496
step: 180, loss: 0.0007403240306302905
step: 190, loss: 0.0005914652138017118
step: 200, loss: 6.75782939651981e-05
step: 210, loss: 0.006513470783829689
step: 220, loss: 0.00010298299457645044
step: 230, loss: 9.66946390690282e-05
step: 240, loss: 0.0009150300757028162
step: 250, loss: 0.0025746189057826996
step: 260, loss: 0.0017730611143633723
step: 270, loss: 0.0027364776469767094
step: 280, loss: 0.00029093792545609176
step: 290, loss: 4.9993948778137565e-05
step: 300, loss: 1.5194605111901183e-05
step: 310, loss: 0.0002776231849566102
step: 320, loss: 2.0464078261284158e-05
step: 330, loss: 0.0013580442173406482
step: 340, loss: 0.0010050712153315544
step: 350, loss: 0.0003335922956466675
step: 360, loss: 0.0013340098084881902
step: 370, loss: 0.0010710321366786957
step: 380, loss: 0.0007561203674413264
step: 390, loss: 0.00044269353384152055
step: 400, loss: 0.0010425535729154944
step: 410, loss: 0.0017983198631554842
step: 420, loss: 0.007050951011478901
step: 430, loss: 0.00987337902188301
step: 440, loss: 0.00931684672832489
step: 450, loss: 2.8179283617646433e-05
step: 460, loss: 0.010247163474559784
step: 470, loss: 4.634244305634638e-06
step: 480, loss: 0.0012928873766213655
step: 490, loss: 0.005279637407511473
step: 500, loss: 0.006765799596905708
step: 510, loss: 0.040243007242679596
step: 520, loss: 0.00016242085257545114
step: 530, loss: 0.007159125525504351
/usr/local/lib/python3.7/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=0.9526462395543175, f1=0.948905109489051,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004819262831006199
step: 10, loss: 0.0025471432600170374
step: 20, loss: 0.0020879157818853855
step: 30, loss: 0.008020568639039993
step: 40, loss: 1.297046856052475e-05
step: 50, loss: 0.0021922134328633547
step: 60, loss: 0.005796506069600582
step: 70, loss: 0.0003122312482446432
step: 80, loss: 0.000302230182569474
step: 90, loss: 0.009818480350077152
step: 100, loss: 5.3510280849877745e-05
step: 110, loss: 0.0015000508865341544
step: 120, loss: 5.263774801278487e-05
step: 130, loss: 0.0002056249650195241
step: 140, loss: 0.0042764111422002316
step: 150, loss: 0.0003001132281497121
step: 160, loss: 9.743366535985842e-05
step: 170, loss: 2.488955760782119e-05
step: 180, loss: 0.03493109345436096
step: 190, loss: 0.0017158306436613202
step: 200, loss: 0.001339470618404448
step: 210, loss: 0.0007789809606038034
step: 220, loss: 1.761884050210938e-05
step: 230, loss: 0.0010833494598045945
step: 240, loss: 7.208247552625835e-05
step: 250, loss: 0.0002889104653149843
step: 260, loss: 2.9916654966655187e-05
step: 270, loss: 0.0007281119469553232
step: 280, loss: 8.95677221706137e-05
step: 290, loss: 0.0013063537189736962
step: 300, loss: 2.1536694475798868e-05
step: 310, loss: 0.045264311134815216
step: 320, loss: 2.137097362719942e-05
step: 330, loss: 0.0002801692171487957
step: 340, loss: 5.99514132773038e-05
step: 350, loss: 0.001502355094999075
step: 360, loss: 9.006182517623529e-05
step: 370, loss: 0.10051305592060089
step: 380, loss: 0.0006785495788790286
step: 390, loss: 0.00013380771270021796
step: 400, loss: 0.004219475202262402
step: 410, loss: 0.0002666882937774062
```



```

step: 420, loss: 0.0001483502855990082
step: 430, loss: 4.057129262946546e-05
step: 440, loss: 0.002719244919717312
step: 450, loss: 0.00516197644174099
step: 460, loss: 0.06513040512800217
step: 470, loss: 0.00012123293709009886
step: 480, loss: 0.0031375782564282417
step: 490, loss: 0.004530301317572594
step: 500, loss: 0.02110281027853489
step: 510, loss: 0.0002836326602846384
step: 520, loss: 3.830116111203097e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 530, loss: 1.4245149941416457e-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=0.9524688509460083, f1=0.9490445859872612,
best_f1=0.9458218549127642
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 [00:27, 206.79it/s]
load_f1 = 0.9547553093259464
real_f1 = 0.9500687127805771
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:25, 171.31it/s]

```

2 BASELINE

2.1 BASELINE STRUCTURED

2.1.1 Beer - Running the matcher

```

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

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

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

```

Selected optimization level 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: 0.4258165955543518
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.32000000000000006, f1=0.27956989247311825,
best_f1=0.27956989247311825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4875902831554413

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.35294117647058826, f1=0.29885057471264365,
best_f1=0.29885057471264365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42487233877182007
/usr/local/lib/python3.7/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=0.3404255319148936, f1=0.2898550724637682,
best_f1=0.29885057471264365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2142159640789032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4150943396226415, f1=0.3142857142857143,
best_f1=0.3142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22363892197608948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5, f1=0.4324324324324324, best_f1=0.4324324324324324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1681942492723465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6111111111111112, f1=0.44897959183673464,
best_f1=0.44897959183673464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3224595785140991
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=0.6470588235294117, f1=0.48888888888888893,
best_f1=0.48888888888888893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22000135481357574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.6153846153846153, f1=0.5714285714285715,
best_f1=0.48888888888888893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31368744373321533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8461538461538461, f1=0.7096774193548386,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06923451274633408

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9285714285714286, f1=0.7096774193548386,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08553913980722427
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8666666666666666, f1=0.7058823529411764,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13705551624298096
/usr/local/lib/python3.7/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=0.9285714285714286, f1=0.7142857142857143,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009924942627549171
/usr/local/lib/python3.7/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=0.9285714285714286, f1=0.689655172413793,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008849281817674637
/usr/local/lib/python3.7/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=0.9285714285714286, f1=0.689655172413793,
best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.014872557483613491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9285714285714286, f1=0.689655172413793,
best_f1=0.7096774193548386
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, 115626.07it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.8571428571428571
real_f1 = 0.8666666666666666
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:25, 173.14it/s]

```

2.1.2 DBLP-ACM - Running the matcher

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

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 437kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 821kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 533kB/s]
Downloading: 100% 501M/501M [00:11<00:00, 45.3MB/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.")
step: 0, loss: 0.5955874919891357
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4690302610397339
step: 20, loss: 0.4641777575016022
step: 30, loss: 0.33971107006073
step: 40, loss: 0.3727409839630127
step: 50, loss: 0.5801430940628052
step: 60, loss: 0.45630067586898804
step: 70, loss: 0.4199107587337494
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.3448459506034851
step: 90, loss: 0.08277036249637604
step: 100, loss: 0.22000078856945038
step: 110, loss: 0.12458377331495285
step: 120, loss: 0.04722507670521736
step: 130, loss: 0.03704921901226044
step: 140, loss: 0.011186089366674423
step: 150, loss: 0.11368035525083542
step: 160, loss: 0.032888878136873245
step: 170, loss: 0.09210441261529922
step: 180, loss: 0.024961018934845924
step: 190, loss: 0.041316039860248566
step: 200, loss: 0.036393534392118454
step: 210, loss: 0.06554815918207169
step: 220, loss: 0.08724705874919891
step: 230, loss: 0.07631669193506241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9603624009060023, f1=0.956221198156682,

```

```

best_f1=0.956221198156682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020556891337037086
step: 10, loss: 0.02633609250187874
step: 20, loss: 0.009903090074658394
step: 30, loss: 0.01116504892706871
step: 40, loss: 0.009992267936468124
step: 50, loss: 0.01957814022898674
step: 60, loss: 0.003416571067646146
step: 70, loss: 0.022078771144151688
step: 80, loss: 0.013870500959455967
step: 90, loss: 0.07190696150064468
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 100, loss: 0.006477453280240297
step: 110, loss: 0.06300286948680878
step: 120, loss: 0.00527158472687006
step: 130, loss: 0.007714841514825821
step: 140, loss: 0.0055678486824035645
step: 150, loss: 0.07547308504581451
step: 160, loss: 0.004120172467082739
step: 170, loss: 0.003688339376822114
step: 180, loss: 0.006056174170225859
step: 190, loss: 0.047307565808296204
step: 200, loss: 0.0061811343766748905
step: 210, loss: 0.003497451078146696
step: 220, loss: 0.002140991622582078
step: 230, loss: 0.0030227121897041798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831649831649831, f1=0.9818181818181818,
best_f1=0.9818181818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004250477999448776
step: 10, loss: 0.0025666416622698307
step: 20, loss: 0.11866859346628189
step: 30, loss: 0.0035621817223727703
step: 40, loss: 0.009539962746202946
step: 50, loss: 0.05648740753531456
step: 60, loss: 0.03926211595535278

```

```

step: 70, loss: 0.08898322284221649
step: 80, loss: 0.04081130772829056
step: 90, loss: 0.012402810156345367
step: 100, loss: 0.0046046725474298
step: 110, loss: 0.016821343451738358
step: 120, loss: 0.001366067910566926
step: 130, loss: 0.11801017820835114
step: 140, loss: 0.004217457491904497
step: 150, loss: 0.027972547337412834
step: 160, loss: 0.0022552404552698135
step: 170, loss: 0.0015904412139207125
step: 180, loss: 0.004086094442754984
step: 190, loss: 0.07696119695901871
step: 200, loss: 0.034618496894836426
step: 210, loss: 0.0026722252368927
step: 220, loss: 0.005938081070780754
step: 230, loss: 0.04460323229432106
/usr/local/lib/python3.7/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=0.9740112994350283, f1=0.9716231555051079,
best_f1=0.9818181818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009669128805398941
step: 10, loss: 0.002863065106794238
step: 20, loss: 0.007024103309959173
step: 30, loss: 0.0034625157713890076
step: 40, loss: 0.07329665869474411
step: 50, loss: 0.026890020817518234
step: 60, loss: 0.004845838528126478
step: 70, loss: 0.01804930903017521
step: 80, loss: 0.0005251768161542714
step: 90, loss: 0.0012663750676438212
step: 100, loss: 0.001076976303011179
step: 110, loss: 0.0005011293687857687
step: 120, loss: 0.0017879616934806108
step: 130, loss: 0.010495166294276714
step: 140, loss: 0.005916884168982506
step: 150, loss: 0.0009500043233856559
step: 160, loss: 0.0012422187719494104
step: 170, loss: 0.007331525441259146
step: 180, loss: 0.12544433772563934
step: 190, loss: 0.01656751148402691
step: 200, loss: 0.062289007008075714
step: 210, loss: 0.0026974196080118418
step: 220, loss: 0.0008044826681725681

```

```

step: 230, loss: 0.04766303300857544
/usr/local/lib/python3.7/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=0.9796380090497738, f1=0.976054732041049,
best_f1=0.9818181818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004120255820453167
step: 10, loss: 0.0024188184179365635
step: 20, loss: 0.0012762685073539615
step: 30, loss: 0.0010821949690580368
step: 40, loss: 0.0013415392022579908
step: 50, loss: 0.05882507935166359
step: 60, loss: 0.04552808403968811
step: 70, loss: 0.013037948869168758
step: 80, loss: 0.01840035617351532
step: 90, loss: 0.04458065330982208
step: 100, loss: 0.0041205910965800285
step: 110, loss: 0.0020006014965474606
step: 120, loss: 0.003956974949687719
step: 130, loss: 0.00037038075970485806
step: 140, loss: 0.004647114779800177
step: 150, loss: 0.024153875187039375
step: 160, loss: 0.00025422521866858006
step: 170, loss: 0.000470235594548285
step: 180, loss: 0.00142327428329736
step: 190, loss: 0.06302881985902786
step: 200, loss: 0.0037392268422991037
step: 210, loss: 0.007678322959691286
step: 220, loss: 0.00044401647755876184
step: 230, loss: 0.0014529322506859899
/usr/local/lib/python3.7/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=0.9831649831649831, f1=0.9829738933030647,
best_f1=0.9818181818181818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002636198769323528
step: 10, loss: 0.0006736709619872272
step: 20, loss: 0.0008040940156206489
step: 30, loss: 0.0010576623026281595
step: 40, loss: 0.0007133882609196007
step: 50, loss: 0.0010999819496646523
step: 60, loss: 0.003875043708831072

```

```

step: 70, loss: 0.008051790297031403
step: 80, loss: 0.001443781889975071
step: 90, loss: 0.05221674591302872
step: 100, loss: 0.0011296496959403157
step: 110, loss: 0.014568598009645939
step: 120, loss: 0.0006068822112865746
step: 130, loss: 0.0011526194866746664
step: 140, loss: 0.003244144609197974
step: 150, loss: 0.0014534574002027512
step: 160, loss: 0.005755376536399126
step: 170, loss: 0.00044770078966394067
step: 180, loss: 0.000655819196254015
step: 190, loss: 0.001220107777044177
step: 200, loss: 0.003739143256098032
step: 210, loss: 0.005477314814925194
step: 220, loss: 0.004475027322769165
step: 230, loss: 0.009604964405298233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9855072463768116, f1=0.9832402234636871,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022684941068291664
step: 10, loss: 0.0007375967688858509
step: 20, loss: 0.0012626335956156254
step: 30, loss: 0.0010408335365355015
step: 40, loss: 0.0016628263983875513
step: 50, loss: 0.0013315826654434204
step: 60, loss: 0.001027556136250496
step: 70, loss: 0.0010919576743617654
step: 80, loss: 0.0073684207163751125
step: 90, loss: 0.010210106149315834
step: 100, loss: 0.000993130262941122
step: 110, loss: 0.000578428793232888
step: 120, loss: 0.0008099551778286695
step: 130, loss: 0.0006245640688575804
step: 140, loss: 0.0005322014912962914
step: 150, loss: 0.009309903718531132
step: 160, loss: 0.0022161260712891817
step: 170, loss: 0.011709081009030342
step: 180, loss: 0.001905698562040925

```

```

step: 190, loss: 0.000981201184913516
step: 200, loss: 0.0020032646134495735
step: 210, loss: 0.0021815034560859203
step: 220, loss: 0.0007062042132019997
step: 230, loss: 0.0033097711857408285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887133182844244, f1=0.9875706214689265,
best_f1=0.9875706214689265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015500533627346158
step: 10, loss: 0.0031693503260612488
step: 20, loss: 0.0020661032758653164
step: 30, loss: 0.0009408447658643126
step: 40, loss: 0.0038266575429588556
step: 50, loss: 0.0008120271377265453
step: 60, loss: 0.0006863222806714475
step: 70, loss: 0.00024036526156123728
step: 80, loss: 0.0553257130086422
step: 90, loss: 0.0004439718904905021
step: 100, loss: 0.0006208589184097946
step: 110, loss: 0.0016806465573608875
step: 120, loss: 0.008634621277451515
step: 130, loss: 0.0086146155372262
step: 140, loss: 0.001143188215792179
step: 150, loss: 0.054679930210113525
step: 160, loss: 0.0016017395537346601
step: 170, loss: 0.057289931923151016
step: 180, loss: 0.0029831579886376858
step: 190, loss: 0.01258473377674818
step: 200, loss: 0.01324030477553606
step: 210, loss: 0.0017379121854901314
step: 220, loss: 0.0018892367370426655
step: 230, loss: 0.0008969053160399199
/usr/local/lib/python3.7/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=0.987736900780379, f1=0.978865406006674,
best_f1=0.9875706214689265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005470059113577008
step: 10, loss: 0.0006026967312209308
step: 20, loss: 0.0029347967356443405
step: 30, loss: 0.002852902514860034
step: 40, loss: 0.0013256387319415808
step: 50, loss: 0.00824160035699606
step: 60, loss: 0.00043902298784814775
step: 70, loss: 0.08480183035135269
step: 80, loss: 0.0004541474627330899
step: 90, loss: 0.023561885580420494
step: 100, loss: 0.000645072665065527
step: 110, loss: 0.000586153706535697
step: 120, loss: 0.0054922811686992645
step: 130, loss: 0.00043028267100453377
step: 140, loss: 0.00055472127860412
step: 150, loss: 0.0024419957771897316
step: 160, loss: 0.003112661885097623
step: 170, loss: 0.00034122669603675604
step: 180, loss: 0.00036973756505176425
step: 190, loss: 0.00018581161566544324
step: 200, loss: 0.00041366767254658043
step: 210, loss: 0.020807737484574318
step: 220, loss: 0.000704303034581244
step: 230, loss: 0.00048410682938992977
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.9864864864864865,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008887479198165238
step: 10, loss: 0.000840427353978157
step: 20, loss: 0.0017382071819156408
step: 30, loss: 0.00013586987915914506
step: 40, loss: 0.004018282983452082
step: 50, loss: 0.0003126604715362191
step: 60, loss: 0.0005065082223154604
step: 70, loss: 0.0009773055789992213
step: 80, loss: 0.0002803603420034051
step: 90, loss: 0.00022722345602232963
step: 100, loss: 0.00017632402887102216

```

```

step: 110, loss: 0.0011769612319767475
step: 120, loss: 0.00015458425332326442
step: 130, loss: 0.0004805699863936752
step: 140, loss: 0.0005340586067177355
step: 150, loss: 0.008368960581719875
step: 160, loss: 7.470963464584202e-05
step: 170, loss: 0.00026404959498904645
step: 180, loss: 0.0009517651633359492
step: 190, loss: 0.00019267127208877355
step: 200, loss: 0.0005301128840073943
step: 210, loss: 0.003474404802545905
step: 220, loss: 0.04213150963187218
step: 230, loss: 0.0003587757528293878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9888641425389755, f1=0.9799554565701558,
best_f1=0.9799554565701558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006252229795791209
step: 10, loss: 0.0005799735081382096
step: 20, loss: 0.00028933637076988816
step: 30, loss: 0.0004273297672625631
step: 40, loss: 6.536702858284116e-05
step: 50, loss: 0.00021776134963147342
step: 60, loss: 0.017357848584651947
step: 70, loss: 0.0011257920414209366
step: 80, loss: 0.0057997927069664
step: 90, loss: 0.0559091717004776
step: 100, loss: 0.0026056719943881035
step: 110, loss: 0.001271202228963375
step: 120, loss: 0.003747781040146947
step: 130, loss: 0.00019656748918350786
step: 140, loss: 0.004176842048764229
step: 150, loss: 0.00015357686788775027
step: 160, loss: 0.010818121954798698
step: 170, loss: 0.007170871365815401
step: 180, loss: 0.001918313791975379
step: 190, loss: 0.0011166558833792806
step: 200, loss: 0.013269763439893723
step: 210, loss: 0.0002727879327721894
step: 220, loss: 0.0011068293824791908

```



```

step: 230, loss: 0.0007349044317379594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898989898989898, f1=0.983277591973244,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036349432775750756
step: 10, loss: 9.869280620478094e-05
step: 20, loss: 0.005836003459990025
step: 30, loss: 0.018559638410806656
step: 40, loss: 0.0010791480308398604
step: 50, loss: 0.002202525269240141
step: 60, loss: 0.0008479990647174418
step: 70, loss: 0.0001634793443372473
step: 80, loss: 6.912565731909126e-05
step: 90, loss: 0.006099258549511433
step: 100, loss: 7.393283885903656e-05
step: 110, loss: 8.08899785624817e-05
step: 120, loss: 8.716970478417352e-05
step: 130, loss: 0.00014701577310916036
step: 140, loss: 0.00018365614232607186
step: 150, loss: 0.00035106041468679905
step: 160, loss: 0.0017551911296322942
step: 170, loss: 0.00045835101627744734
step: 180, loss: 0.00014616631960961968
step: 190, loss: 0.008354021236300468
step: 200, loss: 7.458850450348109e-05
step: 210, loss: 0.0004535147163551301
step: 220, loss: 0.029146188870072365
step: 230, loss: 0.0005622134194709361
/usr/local/lib/python3.7/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=0.9810055865921787,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001808500092010945
step: 10, loss: 0.00096425402443856
step: 20, loss: 0.00015483486640732735

```

```

step: 30, loss: 4.634738797903992e-05
step: 40, loss: 0.0003098914457950741
step: 50, loss: 0.0014160798164084554
step: 60, loss: 0.0007063522934913635
step: 70, loss: 0.0002668453671503812
step: 80, loss: 0.00016013158892747015
step: 90, loss: 8.752385474508628e-05
step: 100, loss: 0.0008512293570674956
step: 110, loss: 0.0010622640838846564
step: 120, loss: 0.0001473517477279529
step: 130, loss: 8.44668538775295e-05
step: 140, loss: 5.902122211409733e-05
step: 150, loss: 9.959270391846076e-05
step: 160, loss: 0.0033956014085561037
step: 170, loss: 6.0139969718875363e-05
step: 180, loss: 0.043594978749752045
step: 190, loss: 9.82835772447288e-05
step: 200, loss: 0.000239019442233257
step: 210, loss: 0.0005006210994906723
step: 220, loss: 7.956984336487949e-05
step: 230, loss: 0.00027952456730417907
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.9831649831649831,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.598820618819445e-05
step: 10, loss: 0.00010614440543577075
step: 20, loss: 0.00013409584062173963
step: 30, loss: 0.0003872533852700144
step: 40, loss: 7.222651765914634e-05
step: 50, loss: 6.393549119820818e-05
step: 60, loss: 4.374955824459903e-05
step: 70, loss: 0.00014369776181410998
step: 80, loss: 8.147869084496051e-05
step: 90, loss: 0.00016148117720149457
step: 100, loss: 0.00010288268094882369
step: 110, loss: 5.8567537053022534e-05
step: 120, loss: 2.0678531654994003e-05
step: 130, loss: 0.00010963861859636381
step: 140, loss: 0.00011387930135242641
step: 150, loss: 2.910745570261497e-05
step: 160, loss: 4.836171865463257e-05
step: 170, loss: 9.170830890070647e-05
step: 180, loss: 5.407094067777507e-05

```

```

step: 190, loss: 4.738889037980698e-05
step: 200, loss: 0.000677631120197475
step: 210, loss: 5.942530697211623e-05
step: 220, loss: 5.4317599278874695e-05
step: 230, loss: 4.1548566514393315e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9899441340782122, f1=0.9789590254706534,
best_f1=0.9789590254706534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011745515279471874
step: 10, loss: 4.2685347580118105e-05
step: 20, loss: 0.0019637688528746367
step: 30, loss: 6.926732021383941e-05
step: 40, loss: 3.484794433461502e-05
step: 50, loss: 3.195797762600705e-05
step: 60, loss: 0.08161837607622147
step: 70, loss: 0.00011168471974087879
step: 80, loss: 4.189577521174215e-05
step: 90, loss: 6.439977005356923e-05
step: 100, loss: 4.593627818394452e-05
step: 110, loss: 0.00019319180864840746
step: 120, loss: 0.0289271492511034
step: 130, loss: 5.570628491113894e-05
step: 140, loss: 0.0054257954470813274
step: 150, loss: 6.241937808226794e-05
step: 160, loss: 0.010261488147079945
step: 170, loss: 4.3685668060788885e-05
step: 180, loss: 8.80536827025935e-05
step: 190, loss: 0.005491448100656271
step: 200, loss: 0.0010489849373698235
step: 210, loss: 0.036418456584215164
step: 220, loss: 7.8697303251829e-05
step: 230, loss: 0.0002826832642313093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9888392857142857, f1=0.9789590254706534,
best_f1=0.9789590254706534
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32

```

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

2473it [00:16, 148.26it/s]

load_f1 = 0.9899441340782122

real_f1 = 0.9899441340782122

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

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

4398it [00:31, 138.74it/s]

2.1.3 DBLP-GoogleScholar - Running the matcher

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

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

step: 0, loss: 0.6393852233886719
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.45871710777282715
step: 20, loss: 0.3413383364677429
step: 30, loss: 0.3538372814655304
step: 40, loss: 0.33377131819725037
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 50, loss: 0.23129233717918396
step: 60, loss: 0.1510351151227951
step: 70, loss: 0.12450778484344482
step: 80, loss: 0.11450627446174622
step: 90, loss: 0.17277966439723969
step: 100, loss: 0.0885126143693924
step: 110, loss: 0.0711611658334732
step: 120, loss: 0.17287978529930115
step: 130, loss: 0.15182678401470184
step: 140, loss: 0.27307480573654175
step: 150, loss: 0.11521054804325104
step: 160, loss: 0.06760808825492859
step: 170, loss: 0.023297283798456192
step: 180, loss: 0.08848536014556885
step: 190, loss: 0.0632375180721283
step: 200, loss: 0.017355045303702354
step: 210, loss: 0.04129467532038689
step: 220, loss: 0.050631675869226456
step: 230, loss: 0.1582104116678238
step: 240, loss: 0.018886404111981392
step: 250, loss: 0.05359029024839401
step: 260, loss: 0.18269577622413635
step: 270, loss: 0.2930072546005249
step: 280, loss: 0.04657242074608803
step: 290, loss: 0.13502007722854614
step: 300, loss: 0.01904098503291607
step: 310, loss: 0.14290766417980194
step: 320, loss: 0.08036140352487564
step: 330, loss: 0.07947714626789093
step: 340, loss: 0.24102698266506195
step: 350, loss: 0.08941333740949631
step: 360, loss: 0.11305331438779831
step: 370, loss: 0.019047554582357407
step: 380, loss: 0.0750788226723671
step: 390, loss: 0.0023001760710030794
step: 400, loss: 0.059490133076906204
step: 410, loss: 0.25732630491256714
step: 420, loss: 0.029567154124379158
step: 430, loss: 0.03213764354586601

```

step: 440, loss: 0.07653414458036423
step: 450, loss: 0.02038203924894333
step: 460, loss: 0.02132071554660797
step: 470, loss: 0.01854737102985382
step: 480, loss: 0.18126481771469116
step: 490, loss: 0.18033716082572937
step: 500, loss: 0.06701326370239258
step: 510, loss: 0.10059145838022232
step: 520, loss: 0.05423269420862198
step: 530, loss: 0.005503544583916664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9390187987161852, f1=0.9376718606782769,
best_f1=0.9376718606782769
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039954256266355515
step: 10, loss: 0.11587103456258774
step: 20, loss: 0.02304951846599579
step: 30, loss: 0.03951505944132805
step: 40, loss: 0.0516178235411644
step: 50, loss: 0.021510986611247063
step: 60, loss: 0.014140921644866467
step: 70, loss: 0.026270687580108643
step: 80, loss: 0.06138930097222328
step: 90, loss: 0.006851667538285255
step: 100, loss: 0.0951671302318573
step: 110, loss: 0.019204361364245415
step: 120, loss: 0.12993451952934265
step: 130, loss: 0.004613752476871014
step: 140, loss: 0.08623816072940826
step: 150, loss: 0.014469613321125507
step: 160, loss: 0.05530261620879173
step: 170, loss: 0.038264717906713486
step: 180, loss: 0.004770695697516203
step: 190, loss: 0.0060236589051783085
step: 200, loss: 0.046573638916015625
step: 210, loss: 0.02111927792429924
step: 220, loss: 0.0014967219904065132
step: 230, loss: 0.0339156836271286
step: 240, loss: 0.05553529039025307
step: 250, loss: 0.06976175308227539

```

```

step: 260, loss: 0.019794225692749023
step: 270, loss: 0.12950348854064941
step: 280, loss: 0.20532304048538208
step: 290, loss: 0.04247497394680977
step: 300, loss: 0.029179979115724564
step: 310, loss: 0.08719193935394287
step: 320, loss: 0.02747999131679535
step: 330, loss: 0.024532360956072807
step: 340, loss: 0.10383787006139755
step: 350, loss: 0.004761488176882267
step: 360, loss: 0.04915276914834976
step: 370, loss: 0.027890756726264954
step: 380, loss: 0.19364336133003235
step: 390, loss: 0.011407413519918919
step: 400, loss: 0.011415408924221992
step: 410, loss: 0.012774641625583172
step: 420, loss: 0.18579673767089844
step: 430, loss: 0.0931035727262497
step: 440, loss: 0.008300362154841423
step: 450, loss: 0.036841992288827896
step: 460, loss: 0.10252797603607178
step: 470, loss: 0.08480022847652435
step: 480, loss: 0.014256716705858707
step: 490, loss: 0.026239069178700447
step: 500, loss: 0.013427411206066608
step: 510, loss: 0.017187276855111122
step: 520, loss: 0.34707826375961304
step: 530, loss: 0.07706505060195923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9452554744525548, f1=0.9432753888380604,
best_f1=0.9432753888380604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12550334632396698
step: 10, loss: 0.08208830654621124
step: 20, loss: 0.011150674894452095
step: 30, loss: 0.0828525498509407
step: 40, loss: 0.039959151297807693
step: 50, loss: 0.002144809812307358
step: 60, loss: 0.00968286395072937
step: 70, loss: 0.018901482224464417

```


step: 80, loss: 0.008479802869260311
step: 90, loss: 0.023887695744633675
step: 100, loss: 0.008877323940396309
step: 110, loss: 0.017304111272096634
step: 120, loss: 0.08271484076976776
step: 130, loss: 0.0320025309920311
step: 140, loss: 0.023076513782143593
step: 150, loss: 0.00947527028620243
step: 160, loss: 0.017557410523295403
step: 170, loss: 0.02061651088297367
step: 180, loss: 0.02405078522861004
step: 190, loss: 0.0025470827240496874
step: 200, loss: 0.013629848137497902
step: 210, loss: 0.041939426213502884
step: 220, loss: 0.05706224963068962
step: 230, loss: 0.026075106114149094
step: 240, loss: 0.0586303249001503
step: 250, loss: 0.037708211690187454
step: 260, loss: 0.06342597305774689
step: 270, loss: 0.03413647785782814
step: 280, loss: 0.0022198217920958996
step: 290, loss: 0.008785508573055267
step: 300, loss: 0.14406388998031616
step: 310, loss: 0.05977155640721321
step: 320, loss: 0.022995179519057274
step: 330, loss: 0.014189926907420158
step: 340, loss: 0.012018095701932907
step: 350, loss: 0.0620930939912796
step: 360, loss: 0.02625163458287716
step: 370, loss: 0.056434035301208496
step: 380, loss: 0.03228013589978218
step: 390, loss: 0.008691211231052876
step: 400, loss: 0.0617409385740757
step: 410, loss: 0.04294430837035179
step: 420, loss: 0.018642781302332878
step: 430, loss: 0.015440122224390507
step: 440, loss: 0.14245043694972992
step: 450, loss: 0.03154245764017105
step: 460, loss: 0.10833264887332916
step: 470, loss: 0.03981318697333336
step: 480, loss: 0.0812232494354248
step: 490, loss: 0.051021602004766464
step: 500, loss: 0.0025609093718230724
step: 510, loss: 0.01504695788025856
step: 520, loss: 0.0019333807285875082
step: 530, loss: 0.024972259998321533
/usr/local/lib/python3.7/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=0.9450757575757576, f1=0.941398865784499,
best_f1=0.9432753888380604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01426005270332098
step: 10, loss: 0.004115786403417587
step: 20, loss: 0.09920191019773483
step: 30, loss: 0.09560286998748779
step: 40, loss: 0.10654884576797485
step: 50, loss: 0.05813773721456528
step: 60, loss: 0.006066151894629002
step: 70, loss: 0.014264695346355438
step: 80, loss: 0.02722550369799137
step: 90, loss: 0.006158432923257351
step: 100, loss: 0.0006275466294027865
step: 110, loss: 0.0022207119036465883
step: 120, loss: 0.009885498322546482
step: 130, loss: 0.10682158917188644
step: 140, loss: 0.024648549035191536
step: 150, loss: 0.00513097969815135
step: 160, loss: 0.004910639487206936
step: 170, loss: 0.007467067334800959
step: 180, loss: 0.08645906299352646
step: 190, loss: 0.030614472925662994
step: 200, loss: 0.01589461974799633
step: 210, loss: 0.050695862621068954
step: 220, loss: 0.024128198623657227
step: 230, loss: 0.019234998151659966
step: 240, loss: 0.027053657919168472
step: 250, loss: 0.08174201846122742
step: 260, loss: 0.00041711816447786987
step: 270, loss: 0.04659423604607582
step: 280, loss: 0.008981358259916306
step: 290, loss: 0.09674791991710663
step: 300, loss: 0.01604365184903145
step: 310, loss: 0.0033579578157514334
step: 320, loss: 0.03938291221857071
step: 330, loss: 0.021108446642756462
step: 340, loss: 0.007703758310526609
step: 350, loss: 0.14098170399665833
step: 360, loss: 0.01264235284179449
step: 370, loss: 0.003259707009419799
step: 380, loss: 0.003012687200680375
step: 390, loss: 0.0009141336777247488
step: 400, loss: 0.018263956531882286
step: 410, loss: 0.002954750554636121
```

```

step: 420, loss: 0.007413115352392197
step: 430, loss: 0.014046445488929749
step: 440, loss: 0.022310996428132057
step: 450, loss: 0.019011706113815308
step: 460, loss: 0.13740412890911102
step: 470, loss: 0.0014830526197329164
step: 480, loss: 0.018457937985658646
step: 490, loss: 0.0008430107263848186
step: 500, loss: 0.010468579828739166
step: 510, loss: 0.06549890339374542
step: 520, loss: 0.007011455949395895
step: 530, loss: 0.0716971904039383
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9567617295308187, f1=0.9458715596330275,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006894845282658935
step: 10, loss: 0.016511667519807816
step: 20, loss: 0.060154251754283905
step: 30, loss: 0.004831334110349417
step: 40, loss: 0.0009520570747554302
step: 50, loss: 0.15346798300743103
step: 60, loss: 0.010821860283613205
step: 70, loss: 0.006570377852767706
step: 80, loss: 0.0015588834648951888
step: 90, loss: 0.13665801286697388
step: 100, loss: 0.10323525220155716
step: 110, loss: 0.002091167727485299
step: 120, loss: 0.21949593722820282
step: 130, loss: 0.010278608649969101
step: 140, loss: 0.014488417655229568
step: 150, loss: 0.016241824254393578
step: 160, loss: 0.005079284776002169
step: 170, loss: 0.09944386780261993
step: 180, loss: 0.021306980401277542
step: 190, loss: 0.008556130342185497
step: 200, loss: 0.0043221116065979
step: 210, loss: 0.0006801725248806179
step: 220, loss: 0.0004178300150670111
step: 230, loss: 0.0005928644677624106

```

```

step: 240, loss: 0.0005327434046193957
step: 250, loss: 0.34817227721214294
step: 260, loss: 0.007692547515034676
step: 270, loss: 0.007448776159435511
step: 280, loss: 0.009665046818554401
step: 290, loss: 0.004932183772325516
step: 300, loss: 0.010179045610129833
step: 310, loss: 0.19069981575012207
step: 320, loss: 0.06876218318939209
step: 330, loss: 0.0033854448702186346
step: 340, loss: 0.03486699238419533
step: 350, loss: 0.0007815364515408874
step: 360, loss: 0.0004174456698819995
step: 370, loss: 0.00022662866103928536
step: 380, loss: 0.00033746781991794705
step: 390, loss: 0.0020274824928492308
step: 400, loss: 0.04337018355727196
step: 410, loss: 0.03337308019399643
step: 420, loss: 0.15771758556365967
step: 430, loss: 0.009160324931144714
step: 440, loss: 0.0023723687045276165
step: 450, loss: 0.05293745920062065
step: 460, loss: 0.057495445013046265
step: 470, loss: 0.1284015029668808
step: 480, loss: 0.015407413244247437
step: 490, loss: 0.032213594764471054
step: 500, loss: 0.005508689675480127
step: 510, loss: 0.000786023389082402
step: 520, loss: 0.057913850992918015
step: 530, loss: 0.08274439722299576
/usr/local/lib/python3.7/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=0.9521152952115295, f1=0.9478628464067637,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018089475110173225
step: 10, loss: 0.006251542828977108
step: 20, loss: 0.05410291254520416
step: 30, loss: 0.0070749069564044476
step: 40, loss: 0.0011192196980118752
step: 50, loss: 0.008687763474881649
step: 60, loss: 0.005299269687384367
step: 70, loss: 0.001070790458470583
step: 80, loss: 0.0008452126057818532
step: 90, loss: 0.03830242529511452

```

```
step: 100, loss: 0.034317124634981155
step: 110, loss: 0.010105184279382229
step: 120, loss: 0.005142183043062687
step: 130, loss: 0.001203284366056323
step: 140, loss: 0.0017951298505067825
step: 150, loss: 0.0006034184480085969
step: 160, loss: 0.06340066343545914
step: 170, loss: 0.004946506582200527
step: 180, loss: 0.014907670207321644
step: 190, loss: 0.0650499165058136
step: 200, loss: 0.015105823054909706
step: 210, loss: 0.004968772642314434
step: 220, loss: 0.0032530678436160088
step: 230, loss: 0.002883440349251032
step: 240, loss: 0.028759203851222992
step: 250, loss: 0.007520767860114574
step: 260, loss: 0.0022132208105176687
step: 270, loss: 0.0016752282390370965
step: 280, loss: 0.030327992513775826
step: 290, loss: 0.0005116831744089723
step: 300, loss: 0.030261371284723282
step: 310, loss: 0.13069309294223785
step: 320, loss: 0.011800955981016159
step: 330, loss: 0.03206872195005417
step: 340, loss: 0.0008759223856031895
step: 350, loss: 0.033877115696668625
step: 360, loss: 0.09607788920402527
step: 370, loss: 0.012422147206962109
step: 380, loss: 0.0014967948663979769
step: 390, loss: 0.0008077439852058887
step: 400, loss: 0.04829525202512741
step: 410, loss: 0.009135022759437561
step: 420, loss: 0.0022809449583292007
step: 430, loss: 9.01270832400769e-05
step: 440, loss: 0.0003495201817713678
step: 450, loss: 0.18234507739543915
step: 460, loss: 0.004111988004297018
step: 470, loss: 0.0023259753361344337
step: 480, loss: 0.0027341400273144245
step: 490, loss: 0.007330968976020813
step: 500, loss: 0.0006666210247203708
step: 510, loss: 0.2626692056655884
step: 520, loss: 0.0010404415661469102
step: 530, loss: 0.013437341898679733
/usr/local/lib/python3.7/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=0.9494855004677268, f1=0.9497206703910613,
```

```
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004799606278538704
step: 10, loss: 0.0017407088307663798
step: 20, loss: 0.0012143098283559084
step: 30, loss: 0.00979443360120058
step: 40, loss: 0.0005806173430755734
step: 50, loss: 0.03741845861077309
step: 60, loss: 0.017140474170446396
step: 70, loss: 0.002465002704411745
step: 80, loss: 0.0005121980211697519
step: 90, loss: 3.0947845516493544e-05
step: 100, loss: 0.0024920469149947166
step: 110, loss: 7.06714199623093e-05
step: 120, loss: 0.0009142194176092744
step: 130, loss: 8.543118019588292e-05
step: 140, loss: 0.005378863774240017
step: 150, loss: 0.0002544877352192998
step: 160, loss: 7.694220403209329e-05
step: 170, loss: 0.0031708506867289543
step: 180, loss: 0.12458009272813797
step: 190, loss: 0.025739135220646858
step: 200, loss: 0.0002601853047963232
step: 210, loss: 0.013893214985728264
step: 220, loss: 0.00014922535046935081
step: 230, loss: 8.905708091333508e-05
step: 240, loss: 0.00512818805873394
step: 250, loss: 0.0031497012823820114
step: 260, loss: 0.005235298071056604
step: 270, loss: 0.00030556536512449384
step: 280, loss: 0.0074585541151463985
step: 290, loss: 0.0023819711059331894
step: 300, loss: 0.0004428380634635687
step: 310, loss: 0.0017090322216972709
step: 320, loss: 0.08059095591306686
step: 330, loss: 4.8480331315658987e-05
step: 340, loss: 0.0014768579276278615
step: 350, loss: 0.0018612714484333992
step: 360, loss: 0.008517489768564701
step: 370, loss: 0.014822135679423809
step: 380, loss: 0.011716840788722038
step: 390, loss: 0.01821344904601574
step: 400, loss: 0.004775132052600384
step: 410, loss: 9.2083471827209e-05
step: 420, loss: 0.11122473329305649
step: 430, loss: 0.0010149317095056176
```

```

step: 440, loss: 0.0010387134971097112
step: 450, loss: 0.0011472913902252913
step: 460, loss: 0.0011045710416510701
step: 470, loss: 0.12077246606349945
step: 480, loss: 0.0040496033616364
step: 490, loss: 0.00014960004773456603
step: 500, loss: 0.0006400662241503596
step: 510, loss: 6.839459820184857e-05
step: 520, loss: 0.0001960444642463699
step: 530, loss: 0.0003286800056230277
/usr/local/lib/python3.7/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=0.9480401093892434, f1=0.949614162505674,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000582836742978543
step: 10, loss: 0.003934652544558048
step: 20, loss: 4.513843305176124e-05
step: 30, loss: 0.00013302102161105722
step: 40, loss: 2.4589493477833457e-05
step: 50, loss: 8.744684600969777e-05
step: 60, loss: 9.684261749498546e-05
step: 70, loss: 0.00022853861446492374
step: 80, loss: 0.01927279308438301
step: 90, loss: 0.0003004594473168254
step: 100, loss: 0.025779075920581818
step: 110, loss: 0.0003199505154043436
step: 120, loss: 0.001809550914913416
step: 130, loss: 0.0005206003552302718
step: 140, loss: 0.06797543913125992
step: 150, loss: 0.0026366664096713066
step: 160, loss: 0.0004233641375321895
step: 170, loss: 0.024718955159187317
step: 180, loss: 0.0387042798101902
step: 190, loss: 0.001198089448735118
step: 200, loss: 0.002403064165264368
step: 210, loss: 0.02318321168422699
step: 220, loss: 0.10134989768266678
step: 230, loss: 0.07347408682107925
step: 240, loss: 0.017091410234570503
step: 250, loss: 0.00022784658358432353
step: 260, loss: 5.482304186443798e-05
step: 270, loss: 0.023037154227495193
step: 280, loss: 0.0001688122865743935
step: 290, loss: 0.0007812575786374509

```

```

step: 300, loss: 4.4771561078960076e-05
step: 310, loss: 0.0019174072658643126
step: 320, loss: 0.000476428511319682
step: 330, loss: 0.014139359816908836
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 340, loss: 0.02515377476811409
step: 350, loss: 0.004158633295446634
step: 360, loss: 0.10361362993717194
step: 370, loss: 0.44539186358451843
step: 380, loss: 0.03819872811436653
step: 390, loss: 0.02614542841911316
step: 400, loss: 0.0019101101206615567
step: 410, loss: 0.0005507659516297281
step: 420, loss: 0.0002087587199639529
step: 430, loss: 0.00036745332181453705
step: 440, loss: 0.0009266132256016135
step: 450, loss: 0.0004101930826436728
step: 460, loss: 0.0020763245411217213
step: 470, loss: 0.13874302804470062
step: 480, loss: 0.0021501025184988976
step: 490, loss: 0.05026206746697426
step: 500, loss: 0.0002275587321491912
step: 510, loss: 0.0017839488573372364
step: 520, loss: 0.0002849060401786119
step: 530, loss: 0.006586551200598478
/usr/local/lib/python3.7/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=0.9533239038189533, f1=0.9475673122342938,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00037646963028237224
step: 10, loss: 0.00921136699616909
step: 20, loss: 0.0003442346060182899
step: 30, loss: 0.044105056673288345
step: 40, loss: 0.004437177442014217
step: 50, loss: 0.0003465092449914664
step: 60, loss: 0.001422392320819199
step: 70, loss: 0.0014946432784199715
step: 80, loss: 0.007345316931605339
step: 90, loss: 0.06681756675243378
step: 100, loss: 0.0008430648013018072
step: 110, loss: 0.03372228890657425
step: 120, loss: 0.0004585980495903641
step: 130, loss: 0.00038072638562880456
step: 140, loss: 0.006220564246177673

```



```

step: 150, loss: 0.012889321893453598
step: 160, loss: 0.00026539547252468765
step: 170, loss: 0.0012589282123371959
step: 180, loss: 0.00021555466810241342
step: 190, loss: 0.0024056469555944204
step: 200, loss: 0.00027774504269473255
step: 210, loss: 0.0021075340919196606
step: 220, loss: 0.0003283265687059611
step: 230, loss: 0.0007845634827390313
step: 240, loss: 6.394445517798886e-05
step: 250, loss: 0.00042053681681863964
step: 260, loss: 0.005444168113172054
step: 270, loss: 0.00023943008272908628
step: 280, loss: 0.008555902168154716
step: 290, loss: 0.0002535737003199756
step: 300, loss: 9.397242683917284e-05
step: 310, loss: 0.13901570439338684
step: 320, loss: 0.0001273796515306458
step: 330, loss: 0.0002602300082799047
step: 340, loss: 0.002579162362962961
step: 350, loss: 0.008135637268424034
step: 360, loss: 0.00039405806455761194
step: 370, loss: 0.00016533327288925648
step: 380, loss: 0.0003378003020770848
step: 390, loss: 0.0002870618482120335
step: 400, loss: 0.005734222941100597
step: 410, loss: 0.0005446508293971419
step: 420, loss: 0.0035885763354599476
step: 430, loss: 0.009630450047552586
step: 440, loss: 0.00035705167101696134
step: 450, loss: 0.003599104005843401
step: 460, loss: 0.00031685008434578776
step: 470, loss: 0.00023147324100136757
step: 480, loss: 0.0002489154867362231
step: 490, loss: 0.0005008133593946695
step: 500, loss: 0.00015575271390844136
step: 510, loss: 0.005026591010391712
step: 520, loss: 0.0027876710519194603
step: 530, loss: 0.02684607356786728
/usr/local/lib/python3.7/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=0.9517684887459806, f1=0.9512419503219871,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003115431172773242

```

step: 10, loss: 0.00034603086533024907
step: 20, loss: 5.943283758824691e-05
step: 30, loss: 0.00010651408229023218
step: 40, loss: 0.000123142875963822
step: 50, loss: 0.00011717052257154137
step: 60, loss: 0.0007427202071994543
step: 70, loss: 7.783887849655002e-05
step: 80, loss: 0.00010036958701675758
step: 90, loss: 4.311038355808705e-05
step: 100, loss: 6.342178676277399e-05
step: 110, loss: 0.028077254071831703
step: 120, loss: 5.1888600864913315e-05
step: 130, loss: 9.2546601081267e-05
step: 140, loss: 0.00011016250937245786
step: 150, loss: 0.00010472557914908975
step: 160, loss: 0.21262116730213165
step: 170, loss: 0.00023967577726580203
step: 180, loss: 0.0005374367465265095
step: 190, loss: 0.00039584090700373054
step: 200, loss: 0.00023133025388233364
step: 210, loss: 0.0029492906760424376
step: 220, loss: 0.0004906083340756595
step: 230, loss: 0.00026946025900542736
step: 240, loss: 0.0005674511194229126
step: 250, loss: 0.00011293828720226884
step: 260, loss: 8.376520418096334e-05
step: 270, loss: 0.0001417267048964277
step: 280, loss: 0.015614611096680164
step: 290, loss: 0.00019347487250342965
step: 300, loss: 0.001911377185024321
step: 310, loss: 0.04004118964076042
step: 320, loss: 0.0030571434181183577
step: 330, loss: 0.0004070887225680053
step: 340, loss: 5.638599759549834e-05
step: 350, loss: 0.018732817843556404
step: 360, loss: 8.131441427394748e-05
step: 370, loss: 0.00032578635727986693
step: 380, loss: 0.002247593132779002
step: 390, loss: 0.0010952689917758107
step: 400, loss: 0.005220629740506411
step: 410, loss: 0.0003081916074734181
step: 420, loss: 0.00021851329074706882
step: 430, loss: 0.0005889320163987577
step: 440, loss: 0.0028042795602232218
step: 450, loss: 0.0002860157983377576
step: 460, loss: 7.04593476257287e-05
step: 470, loss: 0.0018336394568905234
step: 480, loss: 0.0004005461814813316

```

step: 490, loss: 0.00018217235628981143
step: 500, loss: 0.00010152104368899018
step: 510, loss: 6.808750913478434e-05
step: 520, loss: 0.005959969013929367
step: 530, loss: 0.01497824490070343
/usr/local/lib/python3.7/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=0.9567219152854511, f1=0.9547134935304992,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02572127804160118
step: 10, loss: 0.0008524603908881545
step: 20, loss: 0.00042771484004333615
step: 30, loss: 9.416484681423753e-05
step: 40, loss: 0.00048195631825365126
step: 50, loss: 7.286202890099958e-05
step: 60, loss: 5.893786874366924e-05
step: 70, loss: 0.00010221922275377437
step: 80, loss: 0.00032938842196017504
step: 90, loss: 0.0011420458322390914
step: 100, loss: 0.0005534531665034592
step: 110, loss: 7.254025695146993e-05
step: 120, loss: 0.0002211887767771259
step: 130, loss: 0.00014112655480857939
step: 140, loss: 4.293692472856492e-05
step: 150, loss: 0.00034683645935729146
step: 160, loss: 0.00013686285819858313
step: 170, loss: 1.8771010218188167e-05
step: 180, loss: 3.14655335387215e-05
step: 190, loss: 0.00013381308235693723
step: 200, loss: 2.7997291908832267e-05
step: 210, loss: 8.088284084806219e-05
step: 220, loss: 0.024062521755695343
step: 230, loss: 1.1063867532357108e-05
step: 240, loss: 0.0029527288861572742
step: 250, loss: 2.8251308322069235e-05
step: 260, loss: 0.004220257978886366
step: 270, loss: 0.0005106928292661905
step: 280, loss: 0.0006539073656313121
step: 290, loss: 6.81742894812487e-05
step: 300, loss: 4.3756805098382756e-05
step: 310, loss: 0.004604739136993885
step: 320, loss: 5.2967789088143036e-05
step: 330, loss: 1.7441210729884915e-05
step: 340, loss: 0.0012854146771132946

```

```

step: 350, loss: 0.0003717057406902313
step: 360, loss: 0.00013897253666073084
step: 370, loss: 0.00039916051900945604
step: 380, loss: 0.0031142050866037607
step: 390, loss: 0.002687429543584585
step: 400, loss: 4.94517880724743e-05
step: 410, loss: 0.00010352389654144645
step: 420, loss: 0.02414967119693756
step: 430, loss: 0.0041947634890675545
step: 440, loss: 0.00018355606880504638
step: 450, loss: 0.0018176486482843757
step: 460, loss: 0.00041365044307895005
step: 470, loss: 0.00017179574933834374
step: 480, loss: 0.00039064980228431523
step: 490, loss: 0.00015089756925590336
step: 500, loss: 0.0013857929734513164
step: 510, loss: 0.00011037083459086716
step: 520, loss: 9.415929525857791e-05
step: 530, loss: 0.0026155540253967047
/usr/local/lib/python3.7/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=0.9543761638733707, f1=0.9551820728291317,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012128234084229916
step: 10, loss: 0.00022444268688559532
step: 20, loss: 0.0025570944417268038
step: 30, loss: 0.0005887630395591259
step: 40, loss: 7.200794789241627e-05
step: 50, loss: 0.05248482897877693
step: 60, loss: 3.9966140320757404e-05
step: 70, loss: 0.02014010399580002
step: 80, loss: 0.05867846682667732
step: 90, loss: 0.00028072608984075487
step: 100, loss: 0.024814235046505928
step: 110, loss: 0.00010603245755191892
step: 120, loss: 5.154606333235279e-05
step: 130, loss: 5.869779488421045e-05
step: 140, loss: 4.76104105473496e-05
step: 150, loss: 2.8387534257490188e-05
step: 160, loss: 0.00024162231420632452
step: 170, loss: 4.0961229387903586e-05
step: 180, loss: 5.009997767047025e-05
step: 190, loss: 0.0012280174996703863
step: 200, loss: 0.0026833969168365

```

```

step: 210, loss: 0.00012028906348859891
step: 220, loss: 0.00013843194756191224
step: 230, loss: 0.00018306390848010778
step: 240, loss: 0.002362601226195693
step: 250, loss: 2.484623109921813e-05
step: 260, loss: 4.036879545310512e-05
step: 270, loss: 0.005925278645008802
step: 280, loss: 0.0002192443935200572
step: 290, loss: 0.0015684866812080145
step: 300, loss: 0.003307133447378874
step: 310, loss: 0.018003085628151894
step: 320, loss: 0.000573617173358798
step: 330, loss: 0.005755649413913488
step: 340, loss: 4.514306419878267e-05
step: 350, loss: 0.03557039797306061
step: 360, loss: 0.0002870369062293321
step: 370, loss: 0.0005274181021377444
step: 380, loss: 5.5302771215792745e-05
step: 390, loss: 0.019550610333681107
step: 400, loss: 2.226779906777665e-05
step: 410, loss: 0.0004008712712675333
step: 420, loss: 0.0013434007996693254
step: 430, loss: 0.00010873114661080763
step: 440, loss: 0.0001675194944255054
step: 450, loss: 0.0017873116303235292
step: 460, loss: 3.102980917901732e-05
step: 470, loss: 0.00791512243449688
step: 480, loss: 0.00010626567382132635
step: 490, loss: 2.601288906589616e-05
step: 500, loss: 7.349738734774292e-05
step: 510, loss: 0.0001232834765687585
step: 520, loss: 0.00036049395566806197
step: 530, loss: 0.001932897255755961
/usr/local/lib/python3.7/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=0.956440281030445, f1=0.9475177304964538,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.347940118052065e-05
step: 10, loss: 9.342050179839134e-05
step: 20, loss: 5.500163752003573e-05
step: 30, loss: 4.474636079976335e-05
step: 40, loss: 0.00015427089238073677
step: 50, loss: 0.0097280815243721
step: 60, loss: 0.00012889114441350102

```

step: 70, loss: 0.004143678117543459
step: 80, loss: 0.00039514509262517095
step: 90, loss: 0.0020481799729168415
step: 100, loss: 6.138846219982952e-05
step: 110, loss: 6.489954103017226e-05
step: 120, loss: 5.427708674687892e-05
step: 130, loss: 6.74592811265029e-05
step: 140, loss: 9.284985571866855e-05
step: 150, loss: 0.0031548815313726664
step: 160, loss: 0.00011271088442299515
step: 170, loss: 0.007218631450086832
step: 180, loss: 0.0003158251929562539
step: 190, loss: 0.00045826658606529236
step: 200, loss: 2.261535519210156e-05
step: 210, loss: 2.64741393039003e-05
step: 220, loss: 2.5360044674016535e-05
step: 230, loss: 0.003024669364094734
step: 240, loss: 0.0013787730131298304
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 250, loss: 4.77458888781257e-05
step: 260, loss: 0.00025166181148961186
step: 270, loss: 9.962904005078599e-05
step: 280, loss: 2.0741577827720903e-05
step: 290, loss: 5.807812340208329e-05
step: 300, loss: 2.9215749236755073e-05
step: 310, loss: 2.3196298570837826e-05
step: 320, loss: 6.589356053154916e-05
step: 330, loss: 8.030454773688689e-05
step: 340, loss: 0.0022578751668334007
step: 350, loss: 2.2500888007925823e-05
step: 360, loss: 0.07852493971586227
step: 370, loss: 1.4457173165283166e-05
step: 380, loss: 2.5203409677487798e-05
step: 390, loss: 2.9762768463115208e-05
step: 400, loss: 6.930893869139254e-05
step: 410, loss: 3.9370061131194234e-05
step: 420, loss: 0.0006194551824592054
step: 430, loss: 0.00010131680755876005
step: 440, loss: 1.0955793186440133e-05
step: 450, loss: 0.00011293950956314802
step: 460, loss: 0.0009037639829330146
step: 470, loss: 0.001989127602428198
step: 480, loss: 9.081990356207825e-06
step: 490, loss: 0.0007805439527146518
step: 500, loss: 0.0006327811279334128
step: 510, loss: 0.00020464675617404282
step: 520, loss: 3.0249662813730538e-05
step: 530, loss: 2.2775016986997798e-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=0.9566028931404573, f1=0.9539040451552211,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019330058712512255
step: 10, loss: 0.00013752750237472355
step: 20, loss: 1.8249445929541253e-05
step: 30, loss: 0.0002495684602763504
step: 40, loss: 2.7025065719499253e-05
step: 50, loss: 0.00018787347653415054
step: 60, loss: 0.00013394579582381994
step: 70, loss: 5.439109372673556e-05
step: 80, loss: 0.0026804928202182055
step: 90, loss: 3.9748643757775426e-05
step: 100, loss: 1.9382010577828623e-05
step: 110, loss: 0.002386758103966713
step: 120, loss: 2.6361780328443274e-05
step: 130, loss: 1.897609035950154e-05
step: 140, loss: 0.0010342858731746674
step: 150, loss: 0.0001332267711404711
step: 160, loss: 4.8359557695221156e-05
step: 170, loss: 5.983221490168944e-05
step: 180, loss: 7.147128053475171e-05
step: 190, loss: 0.00010121667583007365
step: 200, loss: 3.77962423954159e-05
step: 210, loss: 0.047258585691452026
step: 220, loss: 1.548171348986216e-05
step: 230, loss: 1.7199135982082225e-05
step: 240, loss: 0.0017055124044418335
step: 250, loss: 0.0004008487449027598
step: 260, loss: 8.456972864223644e-05
step: 270, loss: 0.0019490254344418645
step: 280, loss: 0.00021429017942864448
step: 290, loss: 1.2944821719429456e-05
step: 300, loss: 2.6663943572202697e-05
step: 310, loss: 5.537292963708751e-05
step: 320, loss: 2.2172072931425646e-05
step: 330, loss: 6.0076647059759125e-05
step: 340, loss: 7.913713488960639e-05
step: 350, loss: 1.8930702935904264e-05
step: 360, loss: 4.350050585344434e-05
step: 370, loss: 0.00013925778330303729
step: 380, loss: 0.00016000019968487322
step: 390, loss: 0.005704862531274557
```

```

step: 400, loss: 0.0019293763907626271
step: 410, loss: 1.0322496564185712e-05
step: 420, loss: 3.4744567528832704e-05
step: 430, loss: 0.0002513136714696884
step: 440, loss: 0.0005823011742904782
step: 450, loss: 2.5173156245728023e-05
step: 460, loss: 0.03901222348213196
step: 470, loss: 1.520254954812117e-05
step: 480, loss: 2.1609163013636135e-05
step: 490, loss: 0.00016787819913588464
step: 500, loss: 0.0002889079914893955
step: 510, loss: 0.03301278129220009
step: 520, loss: 0.0001809717359719798
step: 530, loss: 0.002243814757093787
/usr/local/lib/python3.7/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=0.953759925268566, f1=0.9549295774647887,
best_f1=0.9458715596330275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.0684850824181922e-05
step: 10, loss: 0.0002981781435664743
step: 20, loss: 0.0017600965220481157
step: 30, loss: 0.011097321286797523
step: 40, loss: 2.8541069696075283e-05
step: 50, loss: 0.0031007234938442707
step: 60, loss: 0.0002637677825987339
step: 70, loss: 1.4457322322414257e-05
step: 80, loss: 1.4453509720624425e-05
step: 90, loss: 2.3563783543067984e-05
step: 100, loss: 6.958691756153712e-06
step: 110, loss: 0.00037485541542991996
step: 120, loss: 1.1887041182490066e-05
step: 130, loss: 0.00020298469462431967
step: 140, loss: 9.777948434930295e-05
step: 150, loss: 1.994064041355159e-05
step: 160, loss: 1.8964396076626144e-05
step: 170, loss: 1.6643114577163942e-05
step: 180, loss: 1.7389122149324976e-05
step: 190, loss: 0.00253259832970798
step: 200, loss: 0.002353334566578269
step: 210, loss: 2.1983805709169246e-05
step: 220, loss: 1.4792714864597656e-05
step: 230, loss: 0.03047296032309532
step: 240, loss: 1.1972785614489112e-05
step: 250, loss: 2.55831764661707e-05

```



```

step: 260, loss: 1.2520330528786872e-05
step: 270, loss: 1.5627185348421335e-05
step: 280, loss: 1.4408778042707127e-05
step: 290, loss: 2.5658700906205922e-05
step: 300, loss: 1.8062799426843412e-05
step: 310, loss: 0.03487420827150345
step: 320, loss: 2.1084179024910554e-05
step: 330, loss: 1.9598133803810924e-05
step: 340, loss: 0.00022505418746732175
step: 350, loss: 1.3537254744733218e-05
step: 360, loss: 1.655091909924522e-05
step: 370, loss: 0.00026750436518341303
step: 380, loss: 3.5807970562018454e-05
step: 390, loss: 4.8726134991738945e-05
step: 400, loss: 0.002008120296522975
step: 410, loss: 4.4884694943903014e-05
step: 420, loss: 1.511281243438134e-05
step: 430, loss: 6.675620625173906e-06
step: 440, loss: 1.8409085896564648e-05
step: 450, loss: 0.006706428714096546
step: 460, loss: 5.608065112028271e-05
step: 470, loss: 1.3958183444628958e-05
step: 480, loss: 0.004676513373851776
step: 490, loss: 0.004011091776192188
step: 500, loss: 0.0011810172582045197
step: 510, loss: 8.955460543802474e-06
step: 520, loss: 5.0981834647245705e-05
step: 530, loss: 7.525000000896398e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9550301344459898, f1=0.9548627268496975,
best_f1=0.9458715596330275
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:35, 161.75it/s]
load_f1 = 0.956884561891516
real_f1 = 0.9565619223659889
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 [00:32, 135.71it/s]

```

2.1.4 Amazon-Google - Running the matcher

```

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

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

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

```

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

step: 0, loss: 0.5241211652755737

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

step: 10, loss: 0.43443790078163147

step: 20, loss: 0.44594648480415344

```

step: 30, loss: 0.33661213517189026
step: 40, loss: 0.41665351390838623
step: 50, loss: 0.45132264494895935
step: 60, loss: 0.45068037509918213
step: 70, loss: 0.2930986285209656
step: 80, loss: 0.3785971701145172
step: 90, loss: 0.30979102849960327
step: 100, loss: 0.22987866401672363
step: 110, loss: 0.24449662864208221
step: 120, loss: 0.3515114486217499
step: 130, loss: 0.3043532371520996
step: 140, loss: 0.4267582893371582
step: 150, loss: 0.30080467462539673
step: 160, loss: 0.37002962827682495
step: 170, loss: 0.21075378358364105
step: 180, loss: 0.15815185010433197
step: 190, loss: 0.38318321108818054
step: 200, loss: 0.19629622995853424
step: 210, loss: 0.3255671560764313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.570342205323194, f1=0.5213270142180094,
best_f1=0.5213270142180094
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2634828984737396
step: 10, loss: 0.10237053781747818
step: 20, loss: 0.32658255100250244
step: 30, loss: 0.44521191716194153
step: 40, loss: 0.4504004716873169
step: 50, loss: 0.138666570186615
step: 60, loss: 0.41935110092163086
step: 70, loss: 0.1288713961839676
step: 80, loss: 0.14023590087890625
step: 90, loss: 0.17284998297691345
step: 100, loss: 0.4762990176677704
step: 110, loss: 0.31396937370300293
step: 120, loss: 0.10989823192358017
step: 130, loss: 0.256140798330307
step: 140, loss: 0.1453375369310379
step: 150, loss: 0.38897812366485596
step: 160, loss: 0.03263266012072563

```

```

step: 170, loss: 0.2489130049943924
step: 180, loss: 0.19307933747768402
step: 190, loss: 0.2904796004295349
step: 200, loss: 0.07206842303276062
step: 210, loss: 0.11232218891382217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6574803149606299, f1=0.6652977412731006,
best_f1=0.6652977412731006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05588917061686516
step: 10, loss: 0.03145894035696983
step: 20, loss: 0.20823484659194946
step: 30, loss: 0.12246821075677872
step: 40, loss: 0.30816149711608887
step: 50, loss: 0.12162040174007416
step: 60, loss: 0.1946038156747818
step: 70, loss: 0.10297788679599762
step: 80, loss: 0.1157074049115181
step: 90, loss: 0.15712690353393555
step: 100, loss: 0.16049812734127045
step: 110, loss: 0.10950452089309692
step: 120, loss: 0.13599929213523865
step: 130, loss: 0.2871009111404419
step: 140, loss: 0.13442067801952362
step: 150, loss: 0.2099481076002121
step: 160, loss: 0.11705181002616882
step: 170, loss: 0.2275076061487198
step: 180, loss: 0.13025183975696564
step: 190, loss: 0.030287671834230423
step: 200, loss: 0.18079069256782532
step: 210, loss: 0.17313237488269806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6783831282952547, f1=0.6678635547576302,
best_f1=0.6678635547576302

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04743769019842148
step: 10, loss: 0.09422306716442108
step: 20, loss: 0.0519057437775192
step: 30, loss: 0.10105524212121964
step: 40, loss: 0.03831828013062477
step: 50, loss: 0.16591061651706696
step: 60, loss: 0.19013521075248718
step: 70, loss: 0.06821215897798538
step: 80, loss: 0.057881806045770645
step: 90, loss: 0.07948202639818192
step: 100, loss: 0.1631375551223755
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.2816023528575897
step: 120, loss: 0.18412256240844727
step: 130, loss: 0.31827449798583984
step: 140, loss: 0.15612748265266418
step: 150, loss: 0.08702666312456131
step: 160, loss: 0.19290360808372498
step: 170, loss: 0.04638296738266945
step: 180, loss: 0.047238729894161224
step: 190, loss: 0.1326654851436615
step: 200, loss: 0.07990981638431549
step: 210, loss: 0.22987307608127594
/usr/local/lib/python3.7/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=0.6554307116104868, f1=0.7033398821218074,
best_f1=0.6678635547576302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13682319223880768
step: 10, loss: 0.041050445288419724
step: 20, loss: 0.08953451365232468
step: 30, loss: 0.005983337294310331
step: 40, loss: 0.051840003579854965
step: 50, loss: 0.13886597752571106
step: 60, loss: 0.09640946984291077
step: 70, loss: 0.06681504845619202
step: 80, loss: 0.06793150305747986
step: 90, loss: 0.0498322956264019
step: 100, loss: 0.014620591886341572
step: 110, loss: 0.1001613512635231
step: 120, loss: 0.06906802207231522
step: 130, loss: 0.08657343685626984

```

```

step: 140, loss: 0.0913141518831253
step: 150, loss: 0.11651775985956192
step: 160, loss: 0.04729800298810005
step: 170, loss: 0.010312489233911037
step: 180, loss: 0.054409824311733246
step: 190, loss: 0.1300656795501709
step: 200, loss: 0.0542294904589653
step: 210, loss: 0.0670556053519249
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6852173913043479, f1=0.6765799256505576,
best_f1=0.6765799256505576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.037920575588941574
step: 10, loss: 0.09935250133275986
step: 20, loss: 0.05047175660729408
step: 30, loss: 0.039044491946697235
step: 40, loss: 0.038307372480630875
step: 50, loss: 0.010777845047414303
step: 60, loss: 0.059045661240816116
step: 70, loss: 0.08215276151895523
step: 80, loss: 0.034306369721889496
step: 90, loss: 0.020267125219106674
step: 100, loss: 0.18734398484230042
step: 110, loss: 0.09078368544578552
step: 120, loss: 0.018923526629805565
step: 130, loss: 0.058542411774396896
step: 140, loss: 0.1004919484257698
step: 150, loss: 0.026655644178390503
step: 160, loss: 0.06387315690517426
step: 170, loss: 0.1156022697687149
step: 180, loss: 0.04351004585623741
step: 190, loss: 0.05037238076329231
step: 200, loss: 0.10671468079090118
step: 210, loss: 0.07325976341962814
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.708955223880597, f1=0.6901960784313725,
best_f1=0.6901960784313725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07021745294332504
step: 10, loss: 0.0025733786169439554
step: 20, loss: 0.011040602810680866
step: 30, loss: 0.12368936836719513
step: 40, loss: 0.02383687160909176
step: 50, loss: 0.035166822373867035
step: 60, loss: 0.03185218945145607
step: 70, loss: 0.026330310851335526
step: 80, loss: 0.02386621944606304
step: 90, loss: 0.10420656949281693
step: 100, loss: 0.027653688564896584
step: 110, loss: 0.1373925507068634
step: 120, loss: 0.02941032126545906
step: 130, loss: 0.02546476386487484
step: 140, loss: 0.015330677852034569
step: 150, loss: 0.039759621024131775
step: 160, loss: 0.3930414915084839
step: 170, loss: 0.027254488319158554
step: 180, loss: 0.0228051096200943
step: 190, loss: 0.0355144627392292
step: 200, loss: 0.017740290611982346
step: 210, loss: 0.1583269089460373
/usr/local/lib/python3.7/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=0.7018255578093306, f1=0.6972860125260961,
best_f1=0.6901960784313725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05609815940260887
step: 10, loss: 0.06729281693696976
step: 20, loss: 0.028711974620819092
step: 30, loss: 0.012478360906243324
step: 40, loss: 0.02371649071574211
step: 50, loss: 0.0010676723904907703
step: 60, loss: 0.010317903943359852
step: 70, loss: 0.03468482941389084
step: 80, loss: 0.08340351283550262
step: 90, loss: 0.07343398034572601
step: 100, loss: 0.09481272846460342
step: 110, loss: 0.02244846522808075

```



```

step: 120, loss: 0.07533644139766693
step: 130, loss: 0.004233808256685734
step: 140, loss: 0.01771051622927189
step: 150, loss: 0.027869828045368195
step: 160, loss: 0.11432760953903198
step: 170, loss: 0.09857817739248276
step: 180, loss: 0.032830387353897095
step: 190, loss: 0.010009906254708767
step: 200, loss: 0.04468733072280884
step: 210, loss: 0.17432977259159088
/usr/local/lib/python3.7/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=0.6956521739130435, f1=0.6824817518248175,
best_f1=0.6901960784313725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026002101600170135
step: 10, loss: 0.04187511280179024
step: 20, loss: 0.03138626739382744
step: 30, loss: 0.0722200945019722
step: 40, loss: 0.018592853099107742
step: 50, loss: 0.016308190301060677
step: 60, loss: 0.00044186145532876253
step: 70, loss: 0.06491202861070633
step: 80, loss: 0.0043890634551644325
step: 90, loss: 0.004557423293590546
step: 100, loss: 0.01568865403532982
step: 110, loss: 0.0914885550737381
step: 120, loss: 0.01677742600440979
step: 130, loss: 0.017997978255152702
step: 140, loss: 0.020551718771457672
step: 150, loss: 0.056913480162620544
step: 160, loss: 0.06069296970963478
step: 170, loss: 0.005469661206007004
step: 180, loss: 0.03341313451528549
step: 190, loss: 0.0033906674943864346
step: 200, loss: 0.04044218361377716
step: 210, loss: 0.025959014892578125
/usr/local/lib/python3.7/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=0.6870229007633588, f1=0.6934865900383141,
best_f1=0.6901960784313725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.003310177708044648
step: 10, loss: 0.006036111619323492
step: 20, loss: 0.0013746211770921946
step: 30, loss: 0.008339562453329563
step: 40, loss: 0.0028481422923505306
step: 50, loss: 0.005405577830970287
step: 60, loss: 0.0022610712330788374
step: 70, loss: 0.004080811515450478
step: 80, loss: 0.013292292132973671
step: 90, loss: 0.010401641950011253
step: 100, loss: 0.14319199323654175
step: 110, loss: 0.0023418564815074205
step: 120, loss: 0.1213766559958458
step: 130, loss: 0.005605265032500029
step: 140, loss: 0.025645600631833076
step: 150, loss: 0.024222280830144882
step: 160, loss: 0.007702468428760767
step: 170, loss: 0.005215305369347334
step: 180, loss: 0.07130532711744308
step: 190, loss: 0.01747550256550312
step: 200, loss: 0.09820973128080368
step: 210, loss: 0.00995662622153759
/usr/local/lib/python3.7/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=0.6971428571428572, f1=0.6937984496124031,
best_f1=0.6901960784313725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05023317039012909
step: 10, loss: 0.011743156239390373
step: 20, loss: 0.08644159138202667
step: 30, loss: 0.018042508512735367
step: 40, loss: 0.018564675003290176
step: 50, loss: 0.0029310635291039944
step: 60, loss: 0.0010050747077912092
step: 70, loss: 0.005632293410599232
step: 80, loss: 0.0022487991955131292
step: 90, loss: 0.1886778175830841
step: 100, loss: 0.037185341119766235
step: 110, loss: 0.05846056714653969
step: 120, loss: 0.022992165759205818
step: 130, loss: 0.003909144084900618
step: 140, loss: 0.022169528529047966
step: 150, loss: 0.013541724532842636
step: 160, loss: 0.0004903601366095245
step: 170, loss: 0.009206129238009453

```

```

step: 180, loss: 0.0699818953871727
step: 190, loss: 0.08750049769878387
step: 200, loss: 0.0040380158461630344
step: 210, loss: 0.019690051674842834
/usr/local/lib/python3.7/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=0.7007575757575758, f1=0.6899224806201552,
best_f1=0.6901960784313725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014514211798086762
step: 10, loss: 0.0007192039047367871
step: 20, loss: 0.04599125310778618
step: 30, loss: 0.0005237089935690165
step: 40, loss: 0.00025256097433157265
step: 50, loss: 0.0014990880154073238
step: 60, loss: 0.01488430891185999
step: 70, loss: 0.00655446108430624
step: 80, loss: 0.1465451717376709
step: 90, loss: 0.026242054998874664
step: 100, loss: 0.007854118943214417
step: 110, loss: 0.0005288972170092165
step: 120, loss: 0.0001578186929691583
step: 130, loss: 0.0017169718630611897
step: 140, loss: 0.0018612346611917019
step: 150, loss: 0.0002283575595356524
step: 160, loss: 0.006571139674633741
step: 170, loss: 0.14552699029445648
step: 180, loss: 0.020484020933508873
step: 190, loss: 0.01714058220386505
step: 200, loss: 0.0027139983139932156
step: 210, loss: 0.0014951807679608464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7142857142857143, f1=0.6980392156862745,
best_f1=0.6980392156862745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0049660042859613895
step: 10, loss: 0.0347110815346241

```

```

step: 20, loss: 0.00040512337000109255
step: 30, loss: 0.0010779881849884987
step: 40, loss: 0.010337267071008682
step: 50, loss: 0.07385195791721344
step: 60, loss: 0.010021871887147427
step: 70, loss: 0.006175587419420481
step: 80, loss: 0.15223127603530884
step: 90, loss: 0.026936467736959457
step: 100, loss: 0.04365905746817589
step: 110, loss: 0.009829921647906303
step: 120, loss: 0.0035276366397738457
step: 130, loss: 0.0003842586593236774
step: 140, loss: 0.017290201038122177
step: 150, loss: 0.001178269274532795
step: 160, loss: 0.05317419022321701
step: 170, loss: 0.013282526284456253
step: 180, loss: 0.000410787935834378
step: 190, loss: 0.0170601774007082
step: 200, loss: 0.05095144733786583
step: 210, loss: 0.002183050150051713
/usr/local/lib/python3.7/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=0.7020408163265306, f1=0.6860706860706861,
best_f1=0.6980392156862745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023832605802454054
step: 10, loss: 0.017756395041942596
step: 20, loss: 0.005341432057321072
step: 30, loss: 0.010725061409175396
step: 40, loss: 0.007356422021985054
step: 50, loss: 0.002100070007145405
step: 60, loss: 0.08556415140628815
step: 70, loss: 0.00029165888554416597
step: 80, loss: 0.04117683321237564
step: 90, loss: 0.0003081090108025819
step: 100, loss: 0.0008806965197436512
step: 110, loss: 0.0004385994398035109
step: 120, loss: 0.00018986270879395306
step: 130, loss: 0.019957704469561577
step: 140, loss: 0.0004471518623176962
step: 150, loss: 0.02866651862859726
step: 160, loss: 0.0009634291636757553
step: 170, loss: 0.05735566094517708
step: 180, loss: 0.0001837316231103614
step: 190, loss: 0.028255503624677658

```

```

step: 200, loss: 0.12247040122747421
step: 210, loss: 0.0014975621597841382
/usr/local/lib/python3.7/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=0.6993603411513859, f1=0.6781857451403888,
best_f1=0.6980392156862745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005461539840325713
step: 10, loss: 0.001192693947814405
step: 20, loss: 0.003064134856685996
step: 30, loss: 0.01980903372168541
step: 40, loss: 0.00041402061469852924
step: 50, loss: 0.00025247156736440957
step: 60, loss: 0.009724412113428116
step: 70, loss: 0.004673611838370562
step: 80, loss: 0.06037049740552902
step: 90, loss: 0.002158720977604389
step: 100, loss: 0.002375432290136814
step: 110, loss: 0.030019087716937065
step: 120, loss: 0.0007589331362396479
step: 130, loss: 8.192969107767567e-05
step: 140, loss: 0.0016218675300478935
step: 150, loss: 0.01599448174238205
step: 160, loss: 0.0008662915206514299
step: 170, loss: 0.005332739092409611
step: 180, loss: 0.009677065536379814
step: 190, loss: 0.0030041190329939127
step: 200, loss: 0.0011892705224454403
step: 210, loss: 0.03724248334765434
/usr/local/lib/python3.7/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=0.6987951807228915, f1=0.6923076923076923,
best_f1=0.6980392156862745
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 [00:09, 234.22it/s]
load_f1 = 0.7065637065637066
real_f1 = 0.7131782945736435
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 133.55it/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 roberta \
    --fp16 \
    --save_model

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

```

Step 3: run the matcher

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

Downloading: 100% 481/481 [00:00<00:00, 438kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 817kB/s]

Downloading: 100% 456k/456k [00:00<00:00, 497kB/s]

Downloading: 100% 501M/501M [00:07<00:00, 62.9MB/s]

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

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

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

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4918926954269409
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4712750315666199
step: 20, loss: 0.2561512291431427
step: 30, loss: 0.35697802901268005
step: 40, loss: 0.2581608295440674
step: 50, loss: 0.30532073974609375
step: 60, loss: 0.4883239269256592
step: 70, loss: 0.4176587760448456
step: 80, loss: 0.16171087324619293
step: 90, loss: 0.3027912676334381
step: 100, loss: 0.43521955609321594
step: 110, loss: 0.23452603816986084
step: 120, loss: 0.34064245223999023
step: 130, loss: 0.3139839172363281
step: 140, loss: 0.1556825190782547
step: 150, loss: 0.2967587411403656
step: 160, loss: 0.18458661437034607
step: 170, loss: 0.5266774892807007
step: 180, loss: 0.11781255155801773
step: 190, loss: 0.10742013901472092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.699228791773779, f1=0.7142857142857143,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3604196310043335
step: 10, loss: 0.3765842616558075
step: 20, loss: 0.6941676139831543
step: 30, loss: 0.1200530081987381
step: 40, loss: 0.24053025245666504
step: 50, loss: 0.2727627456188202
step: 60, loss: 0.3552014231681824
step: 70, loss: 0.17499449849128723
step: 80, loss: 0.08475205302238464
step: 90, loss: 0.13960018754005432
step: 100, loss: 0.16134727001190186

```



```

step: 110, loss: 0.03744382783770561
step: 120, loss: 0.041731663048267365
step: 130, loss: 0.08669155836105347
step: 140, loss: 0.29621821641921997
step: 150, loss: 0.2777673006057739
step: 160, loss: 0.1214185431599617
step: 170, loss: 0.042276762425899506
step: 180, loss: 0.039827801287174225
step: 190, loss: 0.18281544744968414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8134715025906736, f1=0.8142493638676844,
best_f1=0.8142493638676844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08297470211982727
step: 10, loss: 0.11210418492555618
step: 20, loss: 0.01200534775853157
step: 30, loss: 0.024622181430459023
step: 40, loss: 0.028763655573129654
step: 50, loss: 0.06670855730772018
step: 60, loss: 0.09740494191646576
step: 70, loss: 0.05311588570475578
step: 80, loss: 0.2050590068101883
step: 90, loss: 0.0863044261932373
step: 100, loss: 0.04254288226366043
step: 110, loss: 0.19869327545166016
step: 120, loss: 0.16581328213214874
step: 130, loss: 0.12640048563480377
step: 140, loss: 0.05295778065919876
step: 150, loss: 0.08115669339895248
step: 160, loss: 0.142693892121315
step: 170, loss: 0.18727467954158783
step: 180, loss: 0.15022313594818115
step: 190, loss: 0.025480041280388832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.847457627118644, f1=0.848314606741573,
best_f1=0.848314606741573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024139581248164177
step: 10, loss: 0.06780776381492615
step: 20, loss: 0.08978402614593506
step: 30, loss: 0.011460666544735432
step: 40, loss: 0.026381174102425575
step: 50, loss: 0.054863978177309036
step: 60, loss: 0.007984486408531666
step: 70, loss: 0.05622575059533119
step: 80, loss: 0.00821415800601244
step: 90, loss: 0.03466152399778366
step: 100, loss: 0.00904934760183096
step: 110, loss: 0.13390882313251495
step: 120, loss: 0.02936028130352497
step: 130, loss: 0.013858640566468239
step: 140, loss: 0.10921871662139893
step: 150, loss: 0.006890904624015093
step: 160, loss: 0.005702361464500427
step: 170, loss: 0.029777083545923233
step: 180, loss: 0.15494404733181
step: 190, loss: 0.008707121945917606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8507042253521128, f1=0.8547008547008547,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019285976886749268
step: 10, loss: 0.015189510770142078
step: 20, loss: 0.023074183613061905
step: 30, loss: 0.006742666009813547
step: 40, loss: 0.01968725025653839
step: 50, loss: 0.012022417038679123
step: 60, loss: 0.0038200027775019407
step: 70, loss: 0.0030179740861058235
step: 80, loss: 0.010210159234702587
step: 90, loss: 0.0337081141769886
step: 100, loss: 0.0314621776342392

```

```

step: 110, loss: 0.1575438380241394
step: 120, loss: 0.010511073283851147
step: 130, loss: 0.23523326218128204
step: 140, loss: 0.015883704647421837
step: 150, loss: 0.04496880620718002
step: 160, loss: 0.0169965960085392
step: 170, loss: 0.033938731998205185
step: 180, loss: 0.025052174925804138
step: 190, loss: 0.004301917273551226
/usr/local/lib/python3.7/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=0.8254847645429362, f1=0.8459383753501402,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020004378631711006
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.003956093452870846
step: 20, loss: 0.03743424266576767
step: 30, loss: 0.03053566813468933
step: 40, loss: 0.019697025418281555
step: 50, loss: 0.027588501572608948
step: 60, loss: 0.015750855207443237
step: 70, loss: 0.018344448879361153
step: 80, loss: 0.017196182161569595
step: 90, loss: 0.0016079619526863098
step: 100, loss: 0.01484045572578907
step: 110, loss: 0.0680420994758606
step: 120, loss: 0.006002782378345728
step: 130, loss: 0.10363998264074326
step: 140, loss: 0.004328422714024782
step: 150, loss: 0.0027920384891331196
step: 160, loss: 0.015489564277231693
step: 170, loss: 0.44796815514564514
step: 180, loss: 0.01791338063776493
step: 190, loss: 0.009640678763389587
/usr/local/lib/python3.7/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=0.8426966292134831, f1=0.8352272727272728,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007832088507711887
step: 10, loss: 0.0026741756591945887

```

```

step: 20, loss: 0.0012338673695921898
step: 30, loss: 0.017024634405970573
step: 40, loss: 0.0006446252227760851
step: 50, loss: 0.0005442042020149529
step: 60, loss: 0.005648746620863676
step: 70, loss: 0.0011668062070384622
step: 80, loss: 0.0005918241804465652
step: 90, loss: 0.0007380711031146348
step: 100, loss: 0.18310534954071045
step: 110, loss: 0.01162942498922348
step: 120, loss: 0.03183289244771004
step: 130, loss: 0.014619077555835247
step: 140, loss: 0.00565737672150135
step: 150, loss: 0.008044583722949028
step: 160, loss: 0.3963927626609802
step: 170, loss: 0.01907014101743698
step: 180, loss: 0.048856522887945175
step: 190, loss: 0.07643114775419235
/usr/local/lib/python3.7/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=0.842391304347826, f1=0.8429752066115702,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05769599601626396
step: 10, loss: 0.0036962421145290136
step: 20, loss: 0.0009401876013725996
step: 30, loss: 0.12215866893529892
step: 40, loss: 0.021762870252132416
step: 50, loss: 0.024926837533712387
step: 60, loss: 0.011331051588058472
step: 70, loss: 0.0035463757812976837
step: 80, loss: 0.04189949482679367
step: 90, loss: 0.0025916637387126684
step: 100, loss: 0.0011642187600955367
step: 110, loss: 0.02064104750752449
step: 120, loss: 0.019087832421064377
step: 130, loss: 0.0013417252339422703
step: 140, loss: 0.0031616119667887688
step: 150, loss: 0.005573805421590805
step: 160, loss: 0.00317584746517241
step: 170, loss: 0.0016475007869303226
step: 180, loss: 0.008776814676821232
step: 190, loss: 0.000602254644036293
/usr/local/lib/python3.7/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=0.8390501319261214, f1=0.8709677419354839,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007185527356341481
step: 10, loss: 0.00033666682429611683
step: 20, loss: 0.014126715250313282
step: 30, loss: 0.0033174154814332724
step: 40, loss: 0.030074112117290497
step: 50, loss: 0.0058126444928348064
step: 60, loss: 0.004819531459361315
step: 70, loss: 0.0011862042592838407
step: 80, loss: 0.009491339325904846
step: 90, loss: 0.005912358406931162
step: 100, loss: 0.004291401710361242
step: 110, loss: 0.0018359945388510823
step: 120, loss: 0.061788447201251984
step: 130, loss: 0.008493001572787762
step: 140, loss: 0.07157150655984879
step: 150, loss: 0.004947511479258537
step: 160, loss: 0.0018557360162958503
step: 170, loss: 0.013375692069530487
step: 180, loss: 0.0015748664736747742
step: 190, loss: 0.00022076221648603678
/usr/local/lib/python3.7/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=0.8346456692913385, f1=0.8670212765957447,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002555790124461055
step: 10, loss: 0.0006585847586393356
step: 20, loss: 0.0019866004586219788
step: 30, loss: 0.008604292757809162
step: 40, loss: 0.007657147012650967
step: 50, loss: 0.0006530535756610334
step: 60, loss: 0.001117998268455267
step: 70, loss: 0.0007633338682353497
step: 80, loss: 0.001052923733368516
step: 90, loss: 0.0003582282515708357
step: 100, loss: 0.00044309403165243566
step: 110, loss: 0.00029448975692503154
step: 120, loss: 0.0010535630863159895
step: 130, loss: 0.029404444620013237

```

```

step: 140, loss: 0.0003178368497174233
step: 150, loss: 0.00036636582808569074
step: 160, loss: 0.0025085320230573416
step: 170, loss: 0.002864279318600893
step: 180, loss: 0.006748618558049202
step: 190, loss: 0.0004271866928320378
/usr/local/lib/python3.7/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=0.8501362397820164, f1=0.8709677419354839,
best_f1=0.8547008547008547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010061418870463967
step: 10, loss: 0.001045320532284677
step: 20, loss: 0.0005645817145705223
step: 30, loss: 0.038918063044548035
step: 40, loss: 0.00047996515058912337
step: 50, loss: 0.000965704268310219
step: 60, loss: 0.007685643155127764
step: 70, loss: 0.00030771855381317437
step: 80, loss: 0.0013072685105726123
step: 90, loss: 0.00011949526378884912
step: 100, loss: 0.00061986711807549
step: 110, loss: 0.1390155702829361
step: 120, loss: 0.00019956266623921692
step: 130, loss: 0.017027482390403748
step: 140, loss: 0.0013207399751991034
step: 150, loss: 0.0010880683548748493
step: 160, loss: 0.00013807244249619544
step: 170, loss: 0.0004164805286563933
step: 180, loss: 0.007359067909419537
step: 190, loss: 0.0019046271918341517
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8509485094850948, f1=0.854794520547945,
best_f1=0.854794520547945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009981209877878428
step: 10, loss: 0.006849199533462524

```

```

step: 20, loss: 0.012570916675031185
step: 30, loss: 0.09386710077524185
step: 40, loss: 0.00036400946555659175
step: 50, loss: 0.002133840462192893
step: 60, loss: 0.07486637681722641
step: 70, loss: 0.002396195661276579
step: 80, loss: 0.000704685808159411
step: 90, loss: 0.0006146389059722424
step: 100, loss: 0.0008182103629224002
step: 110, loss: 0.0032628776971250772
step: 120, loss: 0.0003571758861653507
step: 130, loss: 0.002284160116687417
step: 140, loss: 0.000169659688253887
step: 150, loss: 0.0012836985988542438
step: 160, loss: 0.00016156755737029016
step: 170, loss: 0.0004095155745744705
step: 180, loss: 0.0001582552067702636
step: 190, loss: 0.0003923574695363641
/usr/local/lib/python3.7/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=0.8435013262599469, f1=0.8586666666666667,
best_f1=0.854794520547945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005571513320319355
step: 10, loss: 0.0004380170430522412
step: 20, loss: 0.00010626792936818674
step: 30, loss: 0.0015533509431406856
step: 40, loss: 0.0003096687432844192
step: 50, loss: 0.00018069504585582763
step: 60, loss: 0.00020842376397922635
step: 70, loss: 0.00024955280241556466
step: 80, loss: 0.00011204877228010446
step: 90, loss: 0.001734739518724382
step: 100, loss: 8.455676288576797e-05
step: 110, loss: 0.00036341301165521145
step: 120, loss: 6.644670065725222e-05
step: 130, loss: 0.00027083922759629786
step: 140, loss: 0.0019030869007110596
step: 150, loss: 0.00015497510321438313
step: 160, loss: 0.0001684119924902916
step: 170, loss: 0.00010332516831113026
step: 180, loss: 0.0001420611224602908
step: 190, loss: 0.11812934279441833
/usr/local/lib/python3.7/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=0.8432432432432432, f1=0.8617886178861788,
best_f1=0.854794520547945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005919057875871658
step: 10, loss: 0.0002581854350864887
step: 20, loss: 0.0005565115134231746
step: 30, loss: 0.003045337973162532
step: 40, loss: 0.00018968958465848118
step: 50, loss: 0.00011139478738186881
step: 60, loss: 0.00021066521003376693
step: 70, loss: 0.0003169781994074583
step: 80, loss: 9.33187548071146e-05
step: 90, loss: 5.671857434208505e-05
step: 100, loss: 0.00021634675795212388
step: 110, loss: 0.00015251168224494904
step: 120, loss: 0.0013881551567465067
step: 130, loss: 0.0007298632990568876
step: 140, loss: 0.0007165850838646293
step: 150, loss: 0.0002918803074862808
step: 160, loss: 0.00010633642523316666
step: 170, loss: 6.758698145858943e-05
step: 180, loss: 0.0007274462841451168
step: 190, loss: 0.0005422031390480697
/usr/local/lib/python3.7/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=0.842391304347826, f1=0.8586956521739131,
best_f1=0.854794520547945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000252085505053401
step: 10, loss: 0.0004639169492293149
step: 20, loss: 0.0011266060173511505
step: 30, loss: 0.002103119622915983
step: 40, loss: 0.0003205986868124455
step: 50, loss: 0.00018508506764192134
step: 60, loss: 0.00013072407455183566
step: 70, loss: 0.00024351802130695432
step: 80, loss: 0.00046377076068893075
step: 90, loss: 0.0006133398856036365
step: 100, loss: 0.001135754631832242
step: 110, loss: 0.0003385645686648786
step: 120, loss: 0.0013206492876634002
step: 130, loss: 0.00018467647896613926

```



```

step: 140, loss: 0.00011075169459218159
step: 150, loss: 8.33084195619449e-05
step: 160, loss: 8.549885387765244e-05
step: 170, loss: 0.0002928194880951196
step: 180, loss: 0.00010179041419178247
step: 190, loss: 0.00010326293704565614
/usr/local/lib/python3.7/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=0.8432432432432432, f1=0.8648648648648648,
best_f1=0.854794520547945
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:13, 148.68it/s]
load_f1 = 0.8548387096774193
real_f1 = 0.8418230563002681
733it [00:00, 3620.52it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 133.76it/s]

```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.49284762144088745
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5136566162109375
step: 20, loss: 0.3395736813545227
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.3686418831348419
step: 40, loss: 0.565153181552887
step: 50, loss: 0.3149006962776184
step: 60, loss: 0.6160929799079895
step: 70, loss: 0.31300657987594604
step: 80, loss: 0.22753013670444489
step: 90, loss: 0.2279149740934372
step: 100, loss: 0.17792610824108124
step: 110, loss: 0.4275267422199249
step: 120, loss: 0.23731297254562378
step: 130, loss: 0.22978562116622925
step: 140, loss: 0.1898873895406723
step: 150, loss: 0.18379954993724823
step: 160, loss: 0.20060628652572632
step: 170, loss: 0.07422297447919846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7531172069825437, f1=0.7737226277372263,
best_f1=0.7737226277372263
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16171807050704956
step: 10, loss: 0.16105054318904877
step: 20, loss: 0.07923094183206558
step: 30, loss: 0.10745285451412201
step: 40, loss: 0.018767962232232094
step: 50, loss: 0.0535123310983181
step: 60, loss: 0.12119244039058685
step: 70, loss: 0.18373382091522217
step: 80, loss: 0.17134533822536469
step: 90, loss: 0.026194889098405838
step: 100, loss: 0.1671024113893509
step: 110, loss: 0.07386984676122665
step: 120, loss: 0.06359045207500458
step: 130, loss: 0.23563185334205627
step: 140, loss: 0.13428816199302673
step: 150, loss: 0.14354108273983002
step: 160, loss: 0.07039191573858261
step: 170, loss: 0.04048002138733864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8243902439024391, f1=0.8253968253968255,
best_f1=0.8253968253968255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1013185903429985
step: 10, loss: 0.09872236847877502
step: 20, loss: 0.08943221718072891
step: 30, loss: 0.05231807753443718
step: 40, loss: 0.13896073400974274
step: 50, loss: 0.0725870355963707
step: 60, loss: 0.025374943390488625
step: 70, loss: 0.17499011754989624
step: 80, loss: 0.041043587028980255
step: 90, loss: 0.19674386084079742
step: 100, loss: 0.0087643563747406
step: 110, loss: 0.10757680982351303
step: 120, loss: 0.14395231008529663

```

```

step: 130, loss: 0.02940693497657776
step: 140, loss: 0.23174616694450378
step: 150, loss: 0.004877994768321514
step: 160, loss: 0.07778275012969971
step: 170, loss: 0.011017462238669395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8312958435207823, f1=0.8454545454545455,
best_f1=0.8454545454545455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020085912197828293
step: 10, loss: 0.019111081957817078
step: 20, loss: 0.0009539254242554307
step: 30, loss: 0.07429079711437225
step: 40, loss: 0.08526033163070679
step: 50, loss: 0.017512142658233643
step: 60, loss: 0.035254258662462234
step: 70, loss: 0.0007419736357405782
step: 80, loss: 0.19125954806804657
step: 90, loss: 0.14047522842884064
step: 100, loss: 0.06961943954229355
step: 110, loss: 0.18598754703998566
step: 120, loss: 0.1354493647813797
step: 130, loss: 0.024766314774751663
step: 140, loss: 0.13865981996059418
step: 150, loss: 0.10321474820375443
step: 160, loss: 0.0006839095731265843
step: 170, loss: 0.057376351207494736
/usr/local/lib/python3.7/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=0.830188679245283, f1=0.8311688311688312,
best_f1=0.8454545454545455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01462436094880104
step: 10, loss: 0.029134521260857582
step: 20, loss: 0.005712189711630344
step: 30, loss: 0.0006824751035310328
step: 40, loss: 0.029590127989649773

```

```

step: 50, loss: 0.004517017863690853
step: 60, loss: 0.06331193447113037
step: 70, loss: 0.004027754999697208
step: 80, loss: 0.0016688397154211998
step: 90, loss: 0.0042341104708611965
step: 100, loss: 0.014366880059242249
step: 110, loss: 0.05879803001880646
step: 120, loss: 0.006094752345234156
step: 130, loss: 0.0004208097525406629
step: 140, loss: 0.08633052557706833
step: 150, loss: 0.042638394981622696
step: 160, loss: 0.003472675569355488
step: 170, loss: 0.02779374271631241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8350515463917526, f1=0.8753056234718827,
best_f1=0.8753056234718827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015072423964738846
step: 10, loss: 0.017319360747933388
step: 20, loss: 0.1253243088722229
step: 30, loss: 0.00844449084252119
step: 40, loss: 0.0015754579799249768
step: 50, loss: 0.004724185448139906
step: 60, loss: 0.011213560588657856
step: 70, loss: 0.0244334377348423
step: 80, loss: 0.0013873694697394967
step: 90, loss: 0.019344763830304146
step: 100, loss: 0.001215797383338213
step: 110, loss: 0.007290862966328859
step: 120, loss: 0.008295570500195026
step: 130, loss: 0.0020958841778337955
step: 140, loss: 0.023816658183932304
step: 150, loss: 0.0021290890872478485
step: 160, loss: 0.2196730375289917
step: 170, loss: 0.024246416985988617
/usr/local/lib/python3.7/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=0.8238095238095239, f1=0.874141876430206,
best_f1=0.8753056234718827

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0324200838804245
step: 10, loss: 0.03538689389824867
step: 20, loss: 0.000934130628593266
step: 30, loss: 0.001382271060720086
step: 40, loss: 0.002989828819409013
step: 50, loss: 0.0007541444501839578
step: 60, loss: 0.007084340788424015
step: 70, loss: 0.005917282775044441
step: 80, loss: 0.016372084617614746
step: 90, loss: 0.03797626122832298
step: 100, loss: 0.00035144377034157515
step: 110, loss: 0.004201752133667469
step: 120, loss: 0.06733693182468414
step: 130, loss: 0.004353213123977184
step: 140, loss: 0.001495116506703198
step: 150, loss: 0.1688673198223114
step: 160, loss: 0.0052458420395851135
step: 170, loss: 0.0458984375
/usr/local/lib/python3.7/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=0.8337730870712401, f1=0.8715365239294711,
best_f1=0.8753056234718827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029308602213859558
step: 10, loss: 0.0015620646299794316
step: 20, loss: 0.0003018030838575214
step: 30, loss: 0.00023605402384418994
step: 40, loss: 0.0002272760175401345
step: 50, loss: 0.03139115124940872
step: 60, loss: 0.0003263259422965348
step: 70, loss: 0.04861677810549736
step: 80, loss: 0.000801338697783649
step: 90, loss: 0.030205806717276573
step: 100, loss: 0.0005440320819616318
step: 110, loss: 0.02606586180627346
step: 120, loss: 0.048637665808200836
step: 130, loss: 0.0008315254235640168
step: 140, loss: 0.027670303359627724
step: 150, loss: 0.0017107619205489755
step: 160, loss: 0.0018502813763916492
step: 170, loss: 0.058890316635370255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8396946564885497, f1=0.8669950738916256,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008108281530439854
step: 10, loss: 0.01551426388323307
step: 20, loss: 0.00750518636777997
step: 30, loss: 0.00043830485083162785
step: 40, loss: 0.043680932372808456
step: 50, loss: 0.000742826028726995
step: 60, loss: 0.010619193315505981
step: 70, loss: 0.0008164162863977253
step: 80, loss: 0.004296497441828251
step: 90, loss: 0.011618314310908318
step: 100, loss: 0.0002320275962119922
step: 110, loss: 0.001307716709561646
step: 120, loss: 0.014925416558980942
step: 130, loss: 0.0002430109743727371
step: 140, loss: 0.000711198546923697
step: 150, loss: 0.05074002593755722
step: 160, loss: 0.10793625563383102
step: 170, loss: 0.052745673805475235
/usr/local/lib/python3.7/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=0.8341968911917099, f1=0.8535353535353535,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003584185615181923
step: 10, loss: 0.0088350223377347
step: 20, loss: 0.0009156054584309459
step: 30, loss: 0.05248911678791046
step: 40, loss: 0.00017958629177883267
step: 50, loss: 0.0035935956984758377
step: 60, loss: 0.006407937966287136
step: 70, loss: 0.0016250057378783822
step: 80, loss: 0.0571129210293293
step: 90, loss: 0.001675884472206235
step: 100, loss: 0.004340074490755796

```



```

step: 110, loss: 0.013525545597076416
step: 120, loss: 0.000524440489243716
step: 130, loss: 0.003596547292545438
step: 140, loss: 0.00016382934700232
step: 150, loss: 0.00012729976151604205
step: 160, loss: 0.00018251432629767805
step: 170, loss: 0.07121390849351883
/usr/local/lib/python3.7/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=0.83248730964467, f1=0.8621553884711779,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032921943347901106
step: 10, loss: 0.01923775114119053
step: 20, loss: 0.00028066744562238455
step: 30, loss: 0.00016752367082517594
step: 40, loss: 0.00010239758557872847
step: 50, loss: 0.01629292592406273
step: 60, loss: 0.007006316911429167
step: 70, loss: 0.0004570115124806762
step: 80, loss: 0.03656401485204697
step: 90, loss: 0.002089596586301923
step: 100, loss: 0.0009314292692579329
step: 110, loss: 8.939770486904308e-05
step: 120, loss: 0.0007657362730242312
step: 130, loss: 0.002084680600091815
step: 140, loss: 0.00014474763884209096
step: 150, loss: 9.934262197930366e-05
step: 160, loss: 0.0013037367025390267
step: 170, loss: 0.001722893095575273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8275862068965518, f1=0.865979381443299,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009823222644627094
step: 10, loss: 0.004458756651729345
step: 20, loss: 7.04884369042702e-05
step: 30, loss: 0.0716029554605484
step: 40, loss: 0.00020296200818847865
step: 50, loss: 0.0014030812308192253
step: 60, loss: 0.00017988981562666595

```

```

step: 70, loss: 4.372554394649342e-05
step: 80, loss: 0.00011753395665436983
step: 90, loss: 0.03998088464140892
step: 100, loss: 4.184467616141774e-05
step: 110, loss: 0.07675523310899734
step: 120, loss: 0.01959347352385521
step: 130, loss: 0.00020462667453102767
step: 140, loss: 0.00016905351367313415
step: 150, loss: 0.019222911447286606
step: 160, loss: 0.0004955215845257044
step: 170, loss: 5.361879084375687e-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=0.8266666666666667, f1=0.865979381443299,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.4515956435352564e-05
step: 10, loss: 7.389934035018086e-05
step: 20, loss: 4.3430682126199827e-05
step: 30, loss: 0.030032511800527573
step: 40, loss: 6.864328315714374e-05
step: 50, loss: 0.00043210037983953953
step: 60, loss: 4.7450041165575385e-05
step: 70, loss: 0.05266920477151871
step: 80, loss: 0.0002953164221253246
step: 90, loss: 5.080265327706002e-05
step: 100, loss: 0.017238717526197433
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.0689442902803421
step: 120, loss: 0.019325880333781242
step: 130, loss: 0.04068581387400627
step: 140, loss: 0.0020209162030369043
step: 150, loss: 0.0008585518808104098
step: 160, loss: 6.2140854424797e-05
step: 170, loss: 0.0005007652216590941
/usr/local/lib/python3.7/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=0.8341968911917099, f1=0.8614609571788413,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008026788011193275
step: 10, loss: 6.861148722236976e-05

```

```

step: 20, loss: 0.008478670381009579
step: 30, loss: 0.001439161249436438
step: 40, loss: 0.18608850240707397
step: 50, loss: 0.00017022198881022632
step: 60, loss: 0.04925263673067093
step: 70, loss: 0.0001756456622388214
step: 80, loss: 0.00019094889285042882
step: 90, loss: 6.971503898967057e-05
step: 100, loss: 0.0005402213428169489
step: 110, loss: 0.01768549531698227
step: 120, loss: 9.980265167541802e-05
step: 130, loss: 7.2010989242699e-05
step: 140, loss: 0.04181574285030365
step: 150, loss: 0.024857331067323685
step: 160, loss: 0.0005649306112900376
step: 170, loss: 0.024471217766404152
/usr/local/lib/python3.7/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=0.8337730870712401, f1=0.8622448979591837,
best_f1=0.8669950738916256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.947024768218398e-05
step: 10, loss: 0.00010217940871370956
step: 20, loss: 0.0004088866990059614
step: 30, loss: 0.00036196320434100926
step: 40, loss: 3.4390985092613846e-05
step: 50, loss: 7.338202703977004e-05
step: 60, loss: 8.402904495596886e-05
step: 70, loss: 0.04973885416984558
step: 80, loss: 0.00015339150559157133
step: 90, loss: 0.010821899399161339
step: 100, loss: 0.04638060927391052
step: 110, loss: 8.099620754364878e-05
step: 120, loss: 7.11971297278069e-05
step: 130, loss: 6.626776303164661e-05
step: 140, loss: 0.013140644878149033
step: 150, loss: 9.910942753776908e-05
step: 160, loss: 3.8458776543848217e-05
step: 170, loss: 9.835214586928487e-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=0.8315789473684211, f1=0.8585858585858587,
best_f1=0.8669950738916256
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, 266.95it/s]
load_f1 = 0.8380462724935733
real_f1 = 0.8396946564885497
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 [00:27, 161.38it/s]
```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

```

--task Dirty/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5864406824111938
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4672294855117798
step: 20, loss: 0.5058885216712952
step: 30, loss: 0.39150887727737427
step: 40, loss: 0.3699277341365814
step: 50, loss: 0.6657893061637878
step: 60, loss: 0.4706275463104248
step: 70, loss: 0.4753074049949646
step: 80, loss: 0.6295117735862732
step: 90, loss: 0.43410322070121765
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.4792053699493408
step: 110, loss: 0.4478342533111572
step: 120, loss: 0.2719002068042755
step: 130, loss: 0.21272537112236023
step: 140, loss: 0.08578447997570038
step: 150, loss: 0.3668353855609894
step: 160, loss: 0.12683098018169403
step: 170, loss: 0.22888582944869995
step: 180, loss: 0.1718243807554245
step: 190, loss: 0.20594905316829681
step: 200, loss: 0.0514933206140995
step: 210, loss: 0.11319036036729813
step: 220, loss: 0.0797996073961258
step: 230, loss: 0.008612542413175106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9391891891891891, f1=0.9399773499433748,
best_f1=0.9399773499433748
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029355892911553383
step: 10, loss: 0.0764969140291214
step: 20, loss: 0.037164319306612015
step: 30, loss: 0.055233750492334366
step: 40, loss: 0.19538797438144684
step: 50, loss: 0.01190901454538107
step: 60, loss: 0.037549685686826706

```

```

step: 70, loss: 0.030061040073633194
step: 80, loss: 0.006378181278705597
step: 90, loss: 0.01223571877926588
step: 100, loss: 0.03025594726204872
step: 110, loss: 0.0842367559671402
step: 120, loss: 0.03906457871198654
step: 130, loss: 0.044689495116472244
step: 140, loss: 0.013020025566220284
step: 150, loss: 0.13051746785640717
step: 160, loss: 0.00977955013513565
step: 170, loss: 0.002861675573512912
step: 180, loss: 0.00348915858194232
step: 190, loss: 0.003688687924295664
step: 200, loss: 0.025760000571608543
step: 210, loss: 0.0753767341375351
step: 220, loss: 0.005262841936200857
step: 230, loss: 0.028020473197102547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9587513935340021, f1=0.9572072072072072,
best_f1=0.9572072072072072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03840611129999161
step: 10, loss: 0.030476843938231468
step: 20, loss: 0.03496444970369339
step: 30, loss: 0.025440895929932594
step: 40, loss: 0.05104538053274155
step: 50, loss: 0.007793285418301821
step: 60, loss: 0.08348522335290909
step: 70, loss: 0.009055613540112972
step: 80, loss: 0.05547572672367096
step: 90, loss: 0.018772760406136513
step: 100, loss: 0.003959941677749157
step: 110, loss: 0.006808184087276459
step: 120, loss: 0.009683715179562569
step: 130, loss: 0.011872156523168087
step: 140, loss: 0.003535831579938531
step: 150, loss: 0.012057634070515633
step: 160, loss: 0.008685552515089512
step: 170, loss: 0.07435259222984314
step: 180, loss: 0.028072498738765717

```

```

step: 190, loss: 0.03628596290946007
step: 200, loss: 0.0795765221118927
step: 210, loss: 0.03838310390710831
step: 220, loss: 0.1460040658712387
step: 230, loss: 0.04138217121362686
/usr/local/lib/python3.7/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=0.9585666293393057, f1=0.9546485260770975,
best_f1=0.9572072072072072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01668514311313629
step: 10, loss: 0.0017161463620141149
step: 20, loss: 0.0016359000001102686
step: 30, loss: 0.0008189461077563465
step: 40, loss: 0.07073606550693512
step: 50, loss: 0.1142001524567604
step: 60, loss: 0.0393516905605793
step: 70, loss: 0.0069188885390758514
step: 80, loss: 0.14781318604946136
step: 90, loss: 0.03759510815143585
step: 100, loss: 0.012678462080657482
step: 110, loss: 0.001296988339163363
step: 120, loss: 0.06564320623874664
step: 130, loss: 0.022543128579854965
step: 140, loss: 0.0060137733817100525
step: 150, loss: 0.0028084951918572187
step: 160, loss: 0.013888712041079998
step: 170, loss: 0.0023217559792101383
step: 180, loss: 0.006110687740147114
step: 190, loss: 0.0011084042489528656
step: 200, loss: 0.15889444947242737
step: 210, loss: 0.005069538950920105
step: 220, loss: 0.0021710048895329237
step: 230, loss: 0.004225531592965126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9722530521642618, f1=0.9663677130044843,
best_f1=0.9663677130044843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005941620096564293
step: 10, loss: 0.00899695884436369
step: 20, loss: 0.12240946292877197
step: 30, loss: 0.02694738656282425
step: 40, loss: 0.05672282353043556
step: 50, loss: 0.10509767383337021
step: 60, loss: 0.04067450761795044
step: 70, loss: 0.010194257833063602
step: 80, loss: 0.030895203351974487
step: 90, loss: 0.03530057519674301
step: 100, loss: 0.0018839577678591013
step: 110, loss: 0.003716943319886923
step: 120, loss: 0.011199990287423134
step: 130, loss: 0.000662054109852761
step: 140, loss: 0.0058610690757632256
step: 150, loss: 0.0078390222042799
step: 160, loss: 0.006523996125906706
step: 170, loss: 0.01029851846396923
step: 180, loss: 0.0025281510315835476
step: 190, loss: 0.1955931931734085
step: 200, loss: 0.04159029200673103
step: 210, loss: 0.005956568289548159
step: 220, loss: 0.00755816325545311
step: 230, loss: 0.01621226966381073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775784753363228, f1=0.9706546275395034,
best_f1=0.9706546275395034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003308493411168456
step: 10, loss: 0.003115664469078183
step: 20, loss: 0.00310559687204659
step: 30, loss: 0.0018954118713736534
step: 40, loss: 0.0009622793295420706
step: 50, loss: 0.002973628230392933
step: 60, loss: 0.003036019392311573
step: 70, loss: 0.1116991639137268
step: 80, loss: 0.003663149196654558
step: 90, loss: 0.0038744774647057056
step: 100, loss: 0.004451482091099024

```

```

step: 110, loss: 0.057972900569438934
step: 120, loss: 0.0013395363930612803
step: 130, loss: 0.0025357494596391916
step: 140, loss: 0.026736196130514145
step: 150, loss: 0.001184971770271659
step: 160, loss: 0.013954016380012035
step: 170, loss: 0.0025075210724025965
step: 180, loss: 0.0019979127682745457
step: 190, loss: 0.0006413131486624479
step: 200, loss: 0.008099248632788658
step: 210, loss: 0.0002730035630520433
step: 220, loss: 0.0172277744859457
step: 230, loss: 0.0006899161962792277
/usr/local/lib/python3.7/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=0.9775784753363228, f1=0.9662162162162162,
best_f1=0.9706546275395034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012253789464011788
step: 10, loss: 0.002193249762058258
step: 20, loss: 0.000821854395326227
step: 30, loss: 0.0004472437431104481
step: 40, loss: 0.0018032253719866276
step: 50, loss: 0.13582047820091248
step: 60, loss: 0.0009459108114242554
step: 70, loss: 0.002882922301068902
step: 80, loss: 0.001549489563331008
step: 90, loss: 0.04952074959874153
step: 100, loss: 0.00043124533840455115
step: 110, loss: 0.0005397950299084187
step: 120, loss: 0.0033026649616658688
step: 130, loss: 0.0020747072994709015
step: 140, loss: 0.00038757783477194607
step: 150, loss: 0.12565280497074127
step: 160, loss: 0.0008905999711714685
step: 170, loss: 0.0005821281229145825
step: 180, loss: 0.0003367135359439999
step: 190, loss: 0.17828933894634247
step: 200, loss: 0.0015234255697578192
step: 210, loss: 0.013381386175751686
step: 220, loss: 0.0013252757489681244
step: 230, loss: 0.003997689578682184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.98, f1=0.9752808988764046, best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014275589492172003
step: 10, loss: 0.00798649899661541
step: 20, loss: 0.003693838370963931
step: 30, loss: 0.0028421450406312943
step: 40, loss: 0.002666605170816183
step: 50, loss: 0.001092863967642188
step: 60, loss: 0.0013354108668863773
step: 70, loss: 0.002832575934007764
step: 80, loss: 0.14459992945194244
step: 90, loss: 0.0007731081568636
step: 100, loss: 0.005177687853574753
step: 110, loss: 0.007411270402371883
step: 120, loss: 0.0006322209374047816
step: 130, loss: 0.005753112956881523
step: 140, loss: 0.0005215908749960363
step: 150, loss: 0.11593642830848694
step: 160, loss: 0.0006744014681316912
step: 170, loss: 0.07998858392238617
step: 180, loss: 0.00038227817276492715
step: 190, loss: 0.004153206944465637
step: 200, loss: 0.010214678943157196
step: 210, loss: 0.0014370951103046536
step: 220, loss: 0.0010730077046900988
step: 230, loss: 0.004440569784492254
/usr/local/lib/python3.7/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=0.9797297297297298, f1=0.9658314350797267,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012624983210116625
step: 10, loss: 0.0007197451777756214
step: 20, loss: 0.0011933278292417526
step: 30, loss: 0.0007577324286103249
step: 40, loss: 0.0016660322435200214
step: 50, loss: 0.0006388345500454307
step: 60, loss: 0.0006573146674782038
step: 70, loss: 0.01390834804624319

```

```

step: 80, loss: 0.0001943305687746033
step: 90, loss: 0.16725124418735504
step: 100, loss: 0.0013234992511570454
step: 110, loss: 0.0013406537473201752
step: 120, loss: 0.05392605811357498
step: 130, loss: 0.008157733827829361
step: 140, loss: 0.0009341894183307886
step: 150, loss: 0.0005864223930984735
step: 160, loss: 0.001342575065791607
step: 170, loss: 0.0004489412240218371
step: 180, loss: 0.0013778316788375378
step: 190, loss: 0.00032786509837023914
step: 200, loss: 0.002974630333483219
step: 210, loss: 0.0786680057644844
step: 220, loss: 0.00031721926643513143
step: 230, loss: 0.0002942800347227603
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843400447427293, f1=0.9752808988764046,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030088244238868356
step: 10, loss: 0.00044713489478453994
step: 20, loss: 0.0003217718913219869
step: 30, loss: 0.00022700949921272695
step: 40, loss: 0.0009226341499015689
step: 50, loss: 0.0002964045852422714
step: 60, loss: 0.0005739384214393795
step: 70, loss: 0.003191860858350992
step: 80, loss: 0.000508254743181169
step: 90, loss: 0.00045945862075313926
step: 100, loss: 0.0004869994299951941
step: 110, loss: 0.0012773983180522919
step: 120, loss: 0.0004052753793075681
step: 130, loss: 0.0049860794097185135
step: 140, loss: 0.000566694070585072
step: 150, loss: 0.0012553422711789608
step: 160, loss: 0.0002300277556059882
step: 170, loss: 0.00037283782148733735
step: 180, loss: 0.0006664919783361256
step: 190, loss: 0.00033402079134248197

```

```

step: 200, loss: 0.0003525417996570468
step: 210, loss: 0.011638772673904896
step: 220, loss: 0.0001933225430548191
step: 230, loss: 0.00019717254326678813
/usr/local/lib/python3.7/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=0.9753363228699552, f1=0.9730337078651685,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013331176887732
step: 10, loss: 0.0003508979280013591
step: 20, loss: 0.00041624062578193843
step: 30, loss: 0.0005227627116255462
step: 40, loss: 9.08122819964774e-05
step: 50, loss: 0.006630332209169865
step: 60, loss: 0.0006954523269087076
step: 70, loss: 0.006077867932617664
step: 80, loss: 0.004651970230042934
step: 90, loss: 0.21448299288749695
step: 100, loss: 0.0005384894902817905
step: 110, loss: 0.00074897485319525
step: 120, loss: 0.0006703276303596795
step: 130, loss: 0.0006128825480118394
step: 140, loss: 0.0004433468566276133
step: 150, loss: 0.0005289417458698153
step: 160, loss: 0.0005528710316866636
step: 170, loss: 0.0005048915045335889
step: 180, loss: 0.0005360572249628603
step: 190, loss: 0.0005074614309705794
step: 200, loss: 0.024933654814958572
step: 210, loss: 0.000377664458937943
step: 220, loss: 0.0005265919608063996
step: 230, loss: 0.0008215924608521163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853768278965129, f1=0.9717514124293786,
best_f1=0.9717514124293786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.000621456594672054
step: 10, loss: 0.00031847009086050093
step: 20, loss: 0.01555243507027626
step: 30, loss: 0.007000530604273081
step: 40, loss: 0.001076473854482174
step: 50, loss: 0.0014049103483557701
step: 60, loss: 0.000585127854719758
step: 70, loss: 0.0006549922400154173
step: 80, loss: 0.000244599039433524
step: 90, loss: 0.004341338761150837
step: 100, loss: 0.0001713510719127953
step: 110, loss: 0.0002187417703680694
step: 120, loss: 0.0005063881981186569
step: 130, loss: 0.0003677325730677694
step: 140, loss: 0.0006086240173317492
step: 150, loss: 0.0008608310017734766
step: 160, loss: 0.0017880230443552136
step: 170, loss: 0.0005387351848185062
step: 180, loss: 0.0006284458213485777
step: 190, loss: 0.0006985347135923803
step: 200, loss: 0.00046804509474895895
step: 210, loss: 0.002485854085534811
step: 220, loss: 0.006448335945606232
step: 230, loss: 0.0003771183255594224
/usr/local/lib/python3.7/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=0.9843400447427293, f1=0.9741863075196409,
best_f1=0.9717514124293786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015941897872835398
step: 10, loss: 0.0004529855796135962
step: 20, loss: 0.00047412904677912593
step: 30, loss: 0.00031630106968805194
step: 40, loss: 0.00045727050746791065
step: 50, loss: 0.0008801344083622098
step: 60, loss: 0.00027543550822883844
step: 70, loss: 0.0003888445207849145
step: 80, loss: 0.0005216528079472482
step: 90, loss: 0.0003531465190462768
step: 100, loss: 0.0005677518784068525
step: 110, loss: 0.0005836496711708605
step: 120, loss: 0.0005030676838941872
step: 130, loss: 0.0005827705026604235
step: 140, loss: 0.0005906478618271649
step: 150, loss: 0.00016550751752220094

```

```

step: 160, loss: 0.003368482692167163
step: 170, loss: 0.0005595298716798425
step: 180, loss: 0.006491715554147959
step: 190, loss: 0.0005921541014686227
step: 200, loss: 0.00014840223593637347
step: 210, loss: 0.0023741889744997025
step: 220, loss: 0.0004636637750081718
step: 230, loss: 0.000378135999199003
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865470852017937, 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: 0.000366062275134027
step: 10, loss: 0.00024179002502933145
step: 20, loss: 0.00035247037885710597
step: 30, loss: 0.0007127332501113415
step: 40, loss: 0.0003793870273511857
step: 50, loss: 0.0003675162442959845
step: 60, loss: 0.00044609475298784673
step: 70, loss: 0.00031894445419311523
step: 80, loss: 0.0007763316389173269
step: 90, loss: 0.0009775320068001747
step: 100, loss: 0.00042196971480734646
step: 110, loss: 0.0006659010541625321
step: 120, loss: 0.00016849164967425168
step: 130, loss: 0.011479475535452366
step: 140, loss: 0.0003785051521845162
step: 150, loss: 0.00023850525030866265
step: 160, loss: 0.0003796009987127036
step: 170, loss: 0.0006492396933026612
step: 180, loss: 0.0004896075115539134
step: 190, loss: 0.0003093595732934773
step: 200, loss: 0.0006502457545138896
step: 210, loss: 0.00034206395503133535
step: 220, loss: 0.000620358856394887
step: 230, loss: 0.003291134024038911
/usr/local/lib/python3.7/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=0.9832026875699889, f1=0.9752808988764046,

```

```

best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006750317988917232
step: 10, loss: 0.0004406913067214191
step: 20, loss: 0.00040032234392128885
step: 30, loss: 0.0005751324351876974
step: 40, loss: 0.00025281813577748835
step: 50, loss: 0.00039422648842446506
step: 60, loss: 0.021477799862623215
step: 70, loss: 0.0005220550228841603
step: 80, loss: 0.0008741935016587377
step: 90, loss: 0.0004057623737026006
step: 100, loss: 0.00026428449200466275
step: 110, loss: 0.0004445348458830267
step: 120, loss: 0.008731343783438206
step: 130, loss: 0.00022761538275517523
step: 140, loss: 0.00302368332631886
step: 150, loss: 0.0007705132011324167
step: 160, loss: 0.0006009471253491938
step: 170, loss: 0.0001703786983853206
step: 180, loss: 0.00047722519957460463
step: 190, loss: 0.0002943655999843031
step: 200, loss: 0.0005854636547155678
step: 210, loss: 0.0005724828224629164
step: 220, loss: 0.00038807577220723033
step: 230, loss: 0.0005725205992348492
/usr/local/lib/python3.7/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=0.9843400447427293, f1=0.9753914988814317,
best_f1=0.9764309764309763
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 [00:14, 176.46it/s]
load_f1 = 0.9853107344632768
real_f1 = 0.9853107344632768
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 [00:26, 163.92it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

```
step: 0, loss: 0.6308673620223999
```

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

```
step: 10, loss: 0.43569323420524597
```

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

```
step: 20, loss: 0.31038400530815125
```

```
step: 30, loss: 0.36797842383384705
```

```
step: 40, loss: 0.19325710833072662
```

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

```
step: 50, loss: 0.21517646312713623
```

```
step: 60, loss: 0.06097866967320442
```

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

```
step: 70, loss: 0.40503713488578796
```

```
step: 80, loss: 0.19431337714195251
```

```
step: 90, loss: 0.12746690213680267
```

```
step: 100, loss: 0.2724410593509674
step: 110, loss: 0.14815251529216766
step: 120, loss: 0.3090895116329193
step: 130, loss: 0.12149252742528915
step: 140, loss: 0.23006704449653625
step: 150, loss: 0.051231611520051956
step: 160, loss: 0.146855428814888
step: 170, loss: 0.06972761452198029
step: 180, loss: 0.09554887562990189
step: 190, loss: 0.02751384675502777
step: 200, loss: 0.055464331060647964
step: 210, loss: 0.04688764363527298
step: 220, loss: 0.005443306174129248
step: 230, loss: 0.29195162653923035
step: 240, loss: 0.029493732377886772
step: 250, loss: 0.049972716718912125
step: 260, loss: 0.07151364535093307
step: 270, loss: 0.4306575059890747
step: 280, loss: 0.047997813671827316
step: 290, loss: 0.02799965813755989
step: 300, loss: 0.025748562067747116
step: 310, loss: 0.07244131714105606
step: 320, loss: 0.04656913876533508
step: 330, loss: 0.18431295454502106
step: 340, loss: 0.4503173232078552
step: 350, loss: 0.10649530589580536
step: 360, loss: 0.13415491580963135
step: 370, loss: 0.12923528254032135
step: 380, loss: 0.274933785200119
step: 390, loss: 0.054872315376996994
step: 400, loss: 0.08395784348249435
step: 410, loss: 0.2393614500761032
step: 420, loss: 0.007686366327106953
step: 430, loss: 0.024318452924489975
step: 440, loss: 0.17436207830905914
step: 450, loss: 0.012102697044610977
step: 460, loss: 0.07848821580410004
step: 470, loss: 0.0769619271159172
step: 480, loss: 0.18159887194633484
step: 490, loss: 0.05256090313196182
step: 500, loss: 0.017573906108736992
step: 510, loss: 0.04050683230161667
step: 520, loss: 0.09293679893016815
step: 530, loss: 0.0033101520966738462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9518916394208314, f1=0.9424964936886395,
best_f1=0.9424964936886395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

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

```
step: 0, loss: 0.0738675594329834
step: 10, loss: 0.05036786571145058
step: 20, loss: 0.00645051384344697
step: 30, loss: 0.0633639544248581
step: 40, loss: 0.060459405183792114
step: 50, loss: 0.04458734765648842
step: 60, loss: 0.04440968111157417
step: 70, loss: 0.016063973307609558
step: 80, loss: 0.09594271332025528
step: 90, loss: 0.009374584071338177
step: 100, loss: 0.1923980563879013
step: 110, loss: 0.016581088304519653
step: 120, loss: 0.02482917159795761
step: 130, loss: 0.0023901090025901794
step: 140, loss: 0.04226371645927429
step: 150, loss: 0.015108383260667324
step: 160, loss: 0.008420602418482304
step: 170, loss: 0.014014222659170628
step: 180, loss: 0.00892513059079647
step: 190, loss: 0.012908535078167915
step: 200, loss: 0.1992916464805603
step: 210, loss: 0.014450805261731148
step: 220, loss: 0.002359607256948948
step: 230, loss: 0.13295944035053253
step: 240, loss: 0.05381322652101517
step: 250, loss: 0.01812053844332695
step: 260, loss: 0.041068170219659805
step: 270, loss: 0.04929543286561966
step: 280, loss: 0.0505107045173645
step: 290, loss: 0.02322838269174099
step: 300, loss: 0.04535858333110809
step: 310, loss: 0.033525533974170685
step: 320, loss: 0.06535109877586365
step: 330, loss: 0.014839465729892254
step: 340, loss: 0.036901552230119705
step: 350, loss: 0.0037129404954612255
step: 360, loss: 0.1259613335132599
step: 370, loss: 0.00282351765781641
step: 380, loss: 0.11557578295469284
step: 390, loss: 0.00666729686781764
```

```

step: 400, loss: 0.1331208199262619
step: 410, loss: 0.015676338225603104
step: 420, loss: 0.007326530292630196
step: 430, loss: 0.18981020152568817
step: 440, loss: 0.012914525344967842
step: 450, loss: 0.020758071914315224
step: 460, loss: 0.04323064163327217
step: 470, loss: 0.02431623637676239
step: 480, loss: 0.0044629801996052265
step: 490, loss: 0.07814677059650421
step: 500, loss: 0.011991379782557487
step: 510, loss: 0.05082060024142265
step: 520, loss: 0.31806400418281555
step: 530, loss: 0.07300733774900436
/usr/local/lib/python3.7/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=0.9516728624535317, f1=0.9481481481481482,
best_f1=0.9424964936886395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1969938725233078
step: 10, loss: 0.01746227592229843
step: 20, loss: 0.008026674389839172
step: 30, loss: 0.041320960968732834
step: 40, loss: 0.08147258311510086
step: 50, loss: 0.004698708187788725
step: 60, loss: 0.00934132281690836
step: 70, loss: 0.0030974787659943104
step: 80, loss: 0.1496574729681015
step: 90, loss: 0.006509755272418261
step: 100, loss: 0.0063647194765508175
step: 110, loss: 0.012028896249830723
step: 120, loss: 0.22453761100769043
step: 130, loss: 0.08099406212568283
step: 140, loss: 0.033413175493478775
step: 150, loss: 0.018241234123706818
step: 160, loss: 0.02718506008386612
step: 170, loss: 0.03737135976552963
step: 180, loss: 0.013694380410015583
step: 190, loss: 0.0099398298189044
step: 200, loss: 0.048155587166547775
step: 210, loss: 0.06689327955245972
step: 220, loss: 0.12564922869205475
step: 230, loss: 0.015880800783634186
step: 240, loss: 0.0898575484752655
step: 250, loss: 0.041469186544418335

```

```

step: 260, loss: 0.16470444202423096
step: 270, loss: 0.002316553145647049
step: 280, loss: 0.009670387953519821
step: 290, loss: 0.008912729099392891
step: 300, loss: 0.18574920296669006
step: 310, loss: 0.12520167231559753
step: 320, loss: 0.018825778737664223
step: 330, loss: 0.007442031987011433
step: 340, loss: 0.006998666562139988
step: 350, loss: 0.11664871871471405
step: 360, loss: 0.015456832945346832
step: 370, loss: 0.023954197764396667
step: 380, loss: 0.007666466757655144
step: 390, loss: 0.004197297152131796
step: 400, loss: 0.14983384311199188
step: 410, loss: 0.01468430645763874
step: 420, loss: 0.0058500515297055244
step: 430, loss: 0.015423450618982315
step: 440, loss: 0.06812790781259537
step: 450, loss: 0.00583431962877512
step: 460, loss: 0.09759047627449036
step: 470, loss: 0.00983015913516283
step: 480, loss: 0.10810495913028717
step: 490, loss: 0.09497496485710144
step: 500, loss: 0.012970101088285446
step: 510, loss: 0.05249495431780815
step: 520, loss: 0.005117934662848711
step: 530, loss: 0.00511776190251112
/usr/local/lib/python3.7/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=0.9487895716945998, f1=0.949767441860465,
best_f1=0.9424964936886395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01209260430186987
step: 10, loss: 0.0014721167972311378
step: 20, loss: 0.0026009015273302794
step: 30, loss: 0.14859648048877716
step: 40, loss: 0.008315942250192165
step: 50, loss: 0.02545805461704731
step: 60, loss: 0.004579576663672924
step: 70, loss: 0.07145723700523376
step: 80, loss: 0.12415937334299088
step: 90, loss: 0.09496353566646576
step: 100, loss: 0.0034639015793800354
step: 110, loss: 0.07776093482971191

```

```
step: 120, loss: 0.008029796183109283
step: 130, loss: 0.025666438043117523
step: 140, loss: 0.0033968379721045494
step: 150, loss: 0.011359786614775658
step: 160, loss: 0.17571556568145752
step: 170, loss: 0.025224804878234863
step: 180, loss: 0.01799466274678707
step: 190, loss: 0.02128271386027336
step: 200, loss: 0.09989558160305023
step: 210, loss: 0.006545179057866335
step: 220, loss: 0.00901031494140625
step: 230, loss: 0.035840462893247604
step: 240, loss: 0.008238530717790127
step: 250, loss: 0.018823275342583656
step: 260, loss: 0.029169170185923576
step: 270, loss: 0.23327863216400146
step: 280, loss: 0.011640479788184166
step: 290, loss: 0.02557649463415146
step: 300, loss: 0.010134106501936913
step: 310, loss: 0.0024708115961402655
step: 320, loss: 0.010847293771803379
step: 330, loss: 0.0168544240295887
step: 340, loss: 0.0015978452283889055
step: 350, loss: 0.01616368442773819
step: 360, loss: 0.06111973896622658
step: 370, loss: 0.0033126724883913994
step: 380, loss: 0.020252499729394913
step: 390, loss: 0.0005427083233371377
step: 400, loss: 0.011656171642243862
step: 410, loss: 0.0385783351957798
step: 420, loss: 0.02702363394200802
step: 430, loss: 0.04048171639442444
step: 440, loss: 0.01890764944255352
step: 450, loss: 0.00567725021392107
step: 460, loss: 0.007562275510281324
step: 470, loss: 0.0016374834813177586
step: 480, loss: 0.02250698022544384
step: 490, loss: 0.005043595097959042
step: 500, loss: 0.008654621429741383
step: 510, loss: 0.007703820243477821
step: 520, loss: 0.04928872361779213
step: 530, loss: 0.14072829484939575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9519365375641624, f1=0.9482517482517483,
best_f1=0.9482517482517483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004916697274893522
step: 10, loss: 0.019446689635515213
step: 20, loss: 0.02122274599969387
step: 30, loss: 0.0020791904535144567
step: 40, loss: 0.0005674568237736821
step: 50, loss: 0.0848168358206749
step: 60, loss: 0.0035145424772053957
step: 70, loss: 0.0009934817207977176
step: 80, loss: 0.003456264268606901
step: 90, loss: 0.11218377202749252
step: 100, loss: 0.01066327653825283
step: 110, loss: 0.010713606141507626
step: 120, loss: 0.07801385223865509
step: 130, loss: 0.004202988930046558
step: 140, loss: 0.0015798771055415273
step: 150, loss: 0.004577525891363621
step: 160, loss: 0.006194422021508217
step: 170, loss: 0.023508165031671524
step: 180, loss: 0.005699211731553078
step: 190, loss: 0.002337457612156868
step: 200, loss: 0.002857899060472846
step: 210, loss: 0.001485330518335104
step: 220, loss: 0.0047107962891459465
step: 230, loss: 0.0030272628646343946
step: 240, loss: 0.0055137332528829575
step: 250, loss: 0.1497051864862442
step: 260, loss: 0.0063744401559233665
step: 270, loss: 0.0021297512575984
step: 280, loss: 0.0037640128284692764
step: 290, loss: 0.08428675681352615
step: 300, loss: 0.11093876510858536
step: 310, loss: 0.054367322474718094
step: 320, loss: 0.006717141717672348
step: 330, loss: 0.00037178193451836705
step: 340, loss: 0.020792139694094658
step: 350, loss: 0.006240156479179859
step: 360, loss: 0.00029532972257584333
step: 370, loss: 0.000612519565038383
step: 380, loss: 0.00027885858435183764
step: 390, loss: 0.021052150055766106
step: 400, loss: 0.000427313701948151
step: 410, loss: 0.16802366077899933
```



```

step: 420, loss: 0.19368864595890045
step: 430, loss: 0.040965672582387924
step: 440, loss: 0.004065255168825388
step: 450, loss: 0.007826893590390682
step: 460, loss: 0.003473056945949793
step: 470, loss: 0.0028377496637403965
step: 480, loss: 0.014983846805989742
step: 490, loss: 0.0033458829857409
step: 500, loss: 0.01566234976053238
step: 510, loss: 0.0008016335195861757
step: 520, loss: 0.08796550333499908
step: 530, loss: 0.007036186754703522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9542056074766355, f1=0.9478098788443615,
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003657157998532057
step: 10, loss: 0.000356143107637763
step: 20, loss: 0.009094388224184513
step: 30, loss: 0.00037679189699701965
step: 40, loss: 0.0015813471982255578
step: 50, loss: 0.000773474806919694
step: 60, loss: 0.02460172399878502
step: 70, loss: 0.0011512935161590576
step: 80, loss: 0.00015917209384497255
step: 90, loss: 0.0013325521722435951
step: 100, loss: 0.0003388152690604329
step: 110, loss: 0.0004036948084831238
step: 120, loss: 0.214355930685997
step: 130, loss: 0.005893392488360405
step: 140, loss: 0.03680785000324249
step: 150, loss: 0.00448536267504096
step: 160, loss: 0.0012102194596081972
step: 170, loss: 0.0020677123684436083
step: 180, loss: 0.001305433688685298
step: 190, loss: 0.09838757663965225
step: 200, loss: 0.02287481166422367
step: 210, loss: 0.00963826384395361
step: 220, loss: 0.0015264999819919467
step: 230, loss: 0.08293716609477997

```

```

step: 240, loss: 0.0037337818648666143
step: 250, loss: 0.004678620025515556
step: 260, loss: 0.0009501404128968716
step: 270, loss: 0.00280697550624609
step: 280, loss: 0.0014769320841878653
step: 290, loss: 0.00023959616373758763
step: 300, loss: 0.027728606015443802
step: 310, loss: 0.004312832839787006
step: 320, loss: 9.664978279033676e-05
step: 330, loss: 0.001434721751138568
step: 340, loss: 0.0004493607266340405
step: 350, loss: 0.0019025527872145176
step: 360, loss: 0.010060708969831467
step: 370, loss: 0.00843252707272768
step: 380, loss: 0.0013709128834307194
step: 390, loss: 0.011005222797393799
step: 400, loss: 0.005657361354678869
step: 410, loss: 0.0055653732270002365
step: 420, loss: 0.0024859735276550055
step: 430, loss: 0.020834337919950485
step: 440, loss: 0.004902938846498728
step: 450, loss: 0.26720482110977173
step: 460, loss: 0.0030119153670966625
step: 470, loss: 0.0022923494689166546
step: 480, loss: 0.004021611530333757
step: 490, loss: 0.006235372740775347
step: 500, loss: 0.001790455775335431
step: 510, loss: 0.01487276703119278
step: 520, loss: 0.000401974655687809
step: 530, loss: 0.000539807544555515
/usr/local/lib/python3.7/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=0.9475620975160993, f1=0.9463056447911886,
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001137280254624784
step: 10, loss: 0.0010780749144032598
step: 20, loss: 0.0077348859049379826
step: 30, loss: 0.004580912180244923
step: 40, loss: 0.0036565798800438643
step: 50, loss: 0.0020408127456903458
step: 60, loss: 0.0007005097577348351
step: 70, loss: 0.0006498401053249836
step: 80, loss: 0.0011583241866901517
step: 90, loss: 0.00023149534536059946

```

```
step: 100, loss: 0.0745650976896286
step: 110, loss: 5.0394712161505595e-05
step: 120, loss: 0.0002558788692113012
step: 130, loss: 5.400013105827384e-05
step: 140, loss: 0.00031956835300661623
step: 150, loss: 0.0004290039069019258
step: 160, loss: 0.00022837256256025285
step: 170, loss: 0.002018974395468831
step: 180, loss: 0.056814443320035934
step: 190, loss: 0.002079145982861519
step: 200, loss: 0.0004078653873875737
step: 210, loss: 0.0010045812232419848
step: 220, loss: 0.0269811749458313
step: 230, loss: 0.002211672253906727
step: 240, loss: 0.010484179481863976
step: 250, loss: 0.003312348620966077
step: 260, loss: 0.001456087571568787
step: 270, loss: 0.0030296901240944862
step: 280, loss: 0.15975072979927063
step: 290, loss: 0.00892794132232666
step: 300, loss: 0.0011195573024451733
step: 310, loss: 0.0008138550911098719
step: 320, loss: 0.001499865553341806
step: 330, loss: 0.006233831401914358
step: 340, loss: 0.004258513916283846
step: 350, loss: 0.0034214917104691267
step: 360, loss: 0.02636951394379139
step: 370, loss: 0.0021061657462269068
step: 380, loss: 0.011744106188416481
step: 390, loss: 0.008245899342000484
step: 400, loss: 0.07183753699064255
step: 410, loss: 0.009181962348520756
step: 420, loss: 0.009563294239342213
step: 430, loss: 0.00099075585603714
step: 440, loss: 0.006052012555301189
step: 450, loss: 0.0015360265970230103
step: 460, loss: 0.003475508652627468
step: 470, loss: 0.14813120663166046
step: 480, loss: 0.002503459807485342
step: 490, loss: 0.0012548526283353567
step: 500, loss: 0.003459368832409382
step: 510, loss: 0.00034298087120987475
step: 520, loss: 0.00024334182671736926
step: 530, loss: 0.00395832397043705
/usr/local/lib/python3.7/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=0.9521597770552717, f1=0.9483480688692415,
```

```
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001447233371436596
step: 10, loss: 0.001269850879907608
step: 20, loss: 0.0002533468941692263
step: 30, loss: 0.000167755366419442
step: 40, loss: 0.00016892075655050576
step: 50, loss: 0.00018623536743689328
step: 60, loss: 0.0015564290806651115
step: 70, loss: 6.464306352427229e-05
step: 80, loss: 0.0012089599622413516
step: 90, loss: 4.889049887424335e-05
step: 100, loss: 8.321314817294478e-05
step: 110, loss: 0.005489380098879337
step: 120, loss: 0.001007336424663663
step: 130, loss: 0.026784269139170647
step: 140, loss: 3.7156129110371694e-05
step: 150, loss: 3.204270979040302e-05
step: 160, loss: 4.9585480155656114e-05
step: 170, loss: 0.004867951851338148
step: 180, loss: 4.482060467125848e-05
step: 190, loss: 0.0005438131047412753
step: 200, loss: 0.002585367998108268
step: 210, loss: 0.05480944365262985
step: 220, loss: 0.0010884612565860152
step: 230, loss: 0.11977387219667435
step: 240, loss: 0.04903800040483475
step: 250, loss: 0.008979900740087032
step: 260, loss: 0.00025016459403559566
step: 270, loss: 0.022727767005562782
step: 280, loss: 0.00022106819960754365
step: 290, loss: 0.011901774443686008
step: 300, loss: 0.00018407529569230974
step: 310, loss: 0.04288572072982788
step: 320, loss: 0.0014662302564829588
step: 330, loss: 0.0002380964724579826
step: 340, loss: 0.151763916015625
step: 350, loss: 6.139175820862874e-05
step: 360, loss: 0.10120277106761932
step: 370, loss: 0.09579125046730042
step: 380, loss: 0.00011133834777865559
step: 390, loss: 0.005405360832810402
step: 400, loss: 0.0121269216760993
step: 410, loss: 0.00036743044620379806
step: 420, loss: 0.0009908305946737528
step: 430, loss: 0.001958580454811454
```

```

step: 440, loss: 0.0028717503882944584
step: 450, loss: 0.0013057797914370894
step: 460, loss: 0.001537178992293775
step: 470, loss: 0.037059154361486435
step: 480, loss: 0.0023323907516896725
step: 490, loss: 0.021672174334526062
step: 500, loss: 0.00010182870755670592
step: 510, loss: 0.0023651698138564825
step: 520, loss: 6.994188152020797e-05
step: 530, loss: 0.0008263495401479304
/usr/local/lib/python3.7/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=0.946124763705104, f1=0.93989588263133,
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0064577399753034115
step: 10, loss: 0.0006118848687037826
step: 20, loss: 0.00011104997247457504
step: 30, loss: 0.006580287124961615
step: 40, loss: 0.0007584752165712416
step: 50, loss: 0.0031687652226537466
step: 60, loss: 0.0013020389014855027
step: 70, loss: 0.1555641144514084
step: 80, loss: 0.0023967125453054905
step: 90, loss: 0.1736822873353958
step: 100, loss: 0.012760825455188751
step: 110, loss: 0.011028088629245758
step: 120, loss: 0.00859018787741661
step: 130, loss: 0.0031699263490736485
step: 140, loss: 0.0002796142944134772
step: 150, loss: 0.0006402014987543225
step: 160, loss: 0.08748070895671844
step: 170, loss: 0.008200306445360184
step: 180, loss: 0.004315913189202547
step: 190, loss: 0.0059159849770367146
step: 200, loss: 0.0006222578813321888
step: 210, loss: 0.0009797952370718122
step: 220, loss: 0.00017380835197400302
step: 230, loss: 0.00024585373466834426
step: 240, loss: 0.0011344046797603369
step: 250, loss: 0.0005374219617806375
step: 260, loss: 0.006231850944459438
step: 270, loss: 0.00031572673469781876
step: 280, loss: 0.0177789144217968
step: 290, loss: 0.0001958147477125749

```

```

step: 300, loss: 0.00015530374366790056
step: 310, loss: 0.01841351017355919
step: 320, loss: 0.006964339409023523
step: 330, loss: 0.00012656519538722932
step: 340, loss: 0.0015568897360935807
step: 350, loss: 0.001259908196516335
step: 360, loss: 9.478615538682789e-05
step: 370, loss: 0.0008822914096526802
step: 380, loss: 0.00020660746667999774
step: 390, loss: 7.464051304850727e-05
step: 400, loss: 0.029060937464237213
step: 410, loss: 0.00013325043255463243
step: 420, loss: 6.200792995514348e-05
step: 430, loss: 0.011351041495800018
step: 440, loss: 0.0012722366955131292
step: 450, loss: 0.0021104076877236366
step: 460, loss: 0.029491418972611427
step: 470, loss: 0.00252247741445899
step: 480, loss: 0.0005065232398919761
step: 490, loss: 0.0191957950592041
step: 500, loss: 8.341016655322164e-05
step: 510, loss: 0.0016530401771888137
step: 520, loss: 0.0020203380845487118
step: 530, loss: 0.00039349248982034624
/usr/local/lib/python3.7/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=0.9520260829063809, f1=0.9475638051044084,
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000232384743867442
step: 10, loss: 0.0009778446983546019
step: 20, loss: 6.400573329301551e-05
step: 30, loss: 0.1347949057817459
step: 40, loss: 5.103125295136124e-05
step: 50, loss: 0.0007414236897602677
step: 60, loss: 0.004144247155636549
step: 70, loss: 7.781755994074047e-05
step: 80, loss: 0.00036734389141201973
step: 90, loss: 3.6783359973924235e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.010577631182968616
step: 110, loss: 0.001716743572615087
step: 120, loss: 1.7764870790415443e-05
step: 130, loss: 0.013953517191112041
step: 140, loss: 0.0003464514738880098

```

```

step: 150, loss: 0.01544242538511753
step: 160, loss: 0.007144998759031296
step: 170, loss: 0.00015214121958706528
step: 180, loss: 0.00013144683907739818
step: 190, loss: 8.354224701179191e-05
step: 200, loss: 0.0001911805447889492
step: 210, loss: 0.00672936299815774
step: 220, loss: 0.0525520034134388
step: 230, loss: 0.00513831153512001
step: 240, loss: 0.0007496240432374179
step: 250, loss: 0.0008378430502489209
step: 260, loss: 0.0007113841129466891
step: 270, loss: 0.00030740309739485383
step: 280, loss: 0.014598353765904903
step: 290, loss: 0.009483563713729382
step: 300, loss: 0.0006392970099113882
step: 310, loss: 0.003294179681688547
step: 320, loss: 0.038833197206258774
step: 330, loss: 0.002236047061160207
step: 340, loss: 0.00024279800709336996
step: 350, loss: 0.0007600177195854485
step: 360, loss: 2.5777932023629546e-05
step: 370, loss: 0.004900245927274227
step: 380, loss: 0.0014093102654442191
step: 390, loss: 8.944082946982235e-05
step: 400, loss: 0.022080153226852417
step: 410, loss: 0.0001337386784143746
step: 420, loss: 3.8174650399014354e-05
step: 430, loss: 0.00017234978440683335
step: 440, loss: 0.006883623544126749
step: 450, loss: 0.00013761752052232623
step: 460, loss: 0.0002417222765628248
step: 470, loss: 0.0010906317038461566
step: 480, loss: 0.00012904875620733947
step: 490, loss: 0.03078516200184822
step: 500, loss: 0.013986513018608093
step: 510, loss: 0.030563656240701675
step: 520, loss: 0.0013254535151645541
step: 530, loss: 0.0008102715946733952
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9552789303826649, f1=0.954001839926403,
best_f1=0.954001839926403

```

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

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

step: 0, loss: 2.3166883693193085e-05
step: 10, loss: 8.089934271993116e-05
step: 20, loss: 3.0678824259666726e-05
step: 30, loss: 2.5412937247892842e-05
step: 40, loss: 3.3327734854537994e-05
step: 50, loss: 6.514543201774359e-05
step: 60, loss: 9.644082456361502e-05
step: 70, loss: 2.6027546482509933e-05
step: 80, loss: 0.00026465795235708356
step: 90, loss: 0.0004823790804948658
step: 100, loss: 0.00015715854533482343
step: 110, loss: 2.787448829622008e-05
step: 120, loss: 3.0549970688298345e-05
step: 130, loss: 2.2339541828841902e-05
step: 140, loss: 2.2879710741108283e-05
step: 150, loss: 8.368113776668906e-05
step: 160, loss: 5.115247768117115e-05
step: 170, loss: 2.408651744190138e-05
step: 180, loss: 1.4912073311279528e-05
step: 190, loss: 9.518402657704428e-05
step: 200, loss: 2.1401083358796313e-05
step: 210, loss: 0.00012639422493521124
step: 220, loss: 0.003022936172783375
step: 230, loss: 0.00010741753067122772
step: 240, loss: 0.0014444788685068488
step: 250, loss: 8.83759930729866e-05
step: 260, loss: 0.0025655822828412056
step: 270, loss: 0.002044414635747671
step: 280, loss: 0.0017095282673835754
step: 290, loss: 2.2451153199654073e-05
step: 300, loss: 0.00011055065260734409
step: 310, loss: 0.00034029193921014667
step: 320, loss: 0.00023461251112166792
step: 330, loss: 1.2881828297395259e-05
step: 340, loss: 9.965147910406813e-05
step: 350, loss: 0.005340413190424442
step: 360, loss: 5.8788242313312367e-05
step: 370, loss: 0.0033893154468387365
step: 380, loss: 0.000664274615701288
step: 390, loss: 0.001312736072577536
step: 400, loss: 6.335452781058848e-05
step: 410, loss: 0.0014519660035148263
step: 420, loss: 8.863040420692414e-05
step: 430, loss: 0.001650368794798851
step: 440, loss: 4.323290704633109e-05


```

step: 450, loss: 0.000900769722647965
step: 460, loss: 0.00021563390328083187
step: 470, loss: 0.0006186337559483945
step: 480, loss: 0.0005702904309146106
step: 490, loss: 3.402995935175568e-05
step: 500, loss: 0.0005883724661543965
step: 510, loss: 8.531938510714099e-05
step: 520, loss: 0.0004733514797408134
step: 530, loss: 0.0001898399059427902
/usr/local/lib/python3.7/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=0.9526044110746129, f1=0.947022972339428,
best_f1=0.954001839926403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.890802640351467e-05
step: 10, loss: 0.000330991402734071
step: 20, loss: 0.0006777825765311718
step: 30, loss: 0.00044123135739937425
step: 40, loss: 0.0009755212813615799
step: 50, loss: 9.361568118038122e-06
step: 60, loss: 1.5076003364811186e-05
step: 70, loss: 8.55972248245962e-05
step: 80, loss: 0.00018311498570255935
step: 90, loss: 0.0002795106265693903
step: 100, loss: 0.010454729199409485
step: 110, loss: 0.00010699187987484038
step: 120, loss: 0.0006403147126547992
step: 130, loss: 0.000853267905768007
step: 140, loss: 0.0004197692614980042
step: 150, loss: 0.000795080151874572
step: 160, loss: 0.00033833985798992217
step: 170, loss: 2.0883149772998877e-05
step: 180, loss: 6.89533626427874e-05
step: 190, loss: 0.00017601842409931123
step: 200, loss: 0.0015529276570305228
step: 210, loss: 8.073101344052702e-05
step: 220, loss: 2.7222315111430362e-05
step: 230, loss: 2.076324562949594e-05
step: 240, loss: 0.001244921120814979
step: 250, loss: 1.026301924866857e-05
step: 260, loss: 1.5694207831984386e-05
step: 270, loss: 0.0015083079924806952
step: 280, loss: 2.1955765987513587e-05
step: 290, loss: 0.003655735868960619
step: 300, loss: 0.0016886297380551696

```

```

step: 310, loss: 4.1681760194478557e-05
step: 320, loss: 1.294889170821989e-05
step: 330, loss: 0.008599316701292992
step: 340, loss: 8.18490661913529e-05
step: 350, loss: 0.000701542419847101
step: 360, loss: 6.587599636986852e-05
step: 370, loss: 0.0004419710021466017
step: 380, loss: 0.0005477808299474418
step: 390, loss: 0.015924610197544098
step: 400, loss: 1.605122997716535e-05
step: 410, loss: 0.0002806921547744423
step: 420, loss: 0.0004538317152764648
step: 430, loss: 0.0005064163124188781
step: 440, loss: 0.025324298068881035
step: 450, loss: 0.0010763033060356975
step: 460, loss: 4.9709280574461445e-05
step: 470, loss: 0.0015031594084575772
step: 480, loss: 0.0002321185020264238
step: 490, loss: 6.381536513799801e-05
step: 500, loss: 2.286247217853088e-05
step: 510, loss: 0.0011744432849809527
step: 520, loss: 0.002464430406689644
step: 530, loss: 6.664067041128874e-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=0.9522041763341067, f1=0.9540441176470589,
best_f1=0.954001839926403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.2397559657983948e-05
step: 10, loss: 1.2658157174882945e-05
step: 20, loss: 6.428502092603594e-05
step: 30, loss: 8.683564374223351e-06
step: 40, loss: 2.6881774829234928e-05
step: 50, loss: 0.0018193774158135056
step: 60, loss: 1.2628371223399881e-05
step: 70, loss: 6.366005982272327e-05
step: 80, loss: 2.0432124074432068e-05
step: 90, loss: 1.9396840798435733e-05
step: 100, loss: 9.667024642112665e-06
step: 110, loss: 1.0370986274210736e-05
step: 120, loss: 9.655846042733174e-06
step: 130, loss: 2.1484116587089375e-05
step: 140, loss: 1.1186905794602353e-05
step: 150, loss: 0.00035840310738421977
step: 160, loss: 0.0004430263361427933

```

```

step: 170, loss: 0.0008699870086275041
step: 180, loss: 1.0941039363387972e-05
step: 190, loss: 2.010073876590468e-05
step: 200, loss: 3.793941141339019e-05
step: 210, loss: 8.445155799563508e-06
step: 220, loss: 8.4935772974859e-06
step: 230, loss: 0.0007196219521574676
step: 240, loss: 0.00038628862239420414
step: 250, loss: 0.0003889111685566604
step: 260, loss: 0.0022256290540099144
step: 270, loss: 0.00042234474676661193
step: 280, loss: 1.084046107280301e-05
step: 290, loss: 8.847452590998728e-06
step: 300, loss: 0.0021009768825024366
step: 310, loss: 3.964909410569817e-05
step: 320, loss: 0.0002921436389442533
step: 330, loss: 1.481152321503032e-05
step: 340, loss: 0.002715568756684661
step: 350, loss: 6.616066002607113e-06
step: 360, loss: 0.09116315096616745
step: 370, loss: 7.551097951363772e-06
step: 380, loss: 1.5921348676783964e-05
step: 390, loss: 7.76717843109509e-06
step: 400, loss: 8.530975173925981e-05
step: 410, loss: 1.5794774299138226e-05
step: 420, loss: 1.160012288892176e-05
step: 430, loss: 0.00048183900071308017
step: 440, loss: 6.77998013998149e-06
step: 450, loss: 0.00045950859203003347
step: 460, loss: 0.001629882724955678
step: 470, loss: 0.0015398190589621663
step: 480, loss: 1.113428697863128e-05
step: 490, loss: 2.384902472840622e-05
step: 500, loss: 8.4600278569269e-06
step: 510, loss: 7.953425665618852e-06
step: 520, loss: 6.485690391855314e-06
step: 530, loss: 2.5203398763551377e-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=0.9542668552569542, f1=0.9454887218045113,
best_f1=0.954001839926403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008040149696171284
step: 10, loss: 0.00021541249589063227
step: 20, loss: 7.852844646549784e-06

```

step: 30, loss: 0.0005984738236293197
step: 40, loss: 0.0005216557765379548
step: 50, loss: 0.002680788515135646
step: 60, loss: 0.01136599201709032
step: 70, loss: 8.720824553165585e-06
step: 80, loss: 1.2006335964542814e-05
step: 90, loss: 6.120619673311012e-06
step: 100, loss: 8.668662303534802e-06
step: 110, loss: 9.581343874742743e-06
step: 120, loss: 8.013036676857155e-06
step: 130, loss: 0.0012040807632729411
step: 140, loss: 0.00038695454713888466
step: 150, loss: 0.0010603846749290824
step: 160, loss: 9.141725968220271e-06
step: 170, loss: 7.968328645802103e-06
step: 180, loss: 1.1655970411084127e-05
step: 190, loss: 0.0002068728645099327
step: 200, loss: 1.0598338121781126e-05
step: 210, loss: 8.873532351572067e-06
step: 220, loss: 8.38555206428282e-06
step: 230, loss: 1.5212859580060467e-05
step: 240, loss: 0.0012172601418569684
step: 250, loss: 0.0009293803013861179
step: 260, loss: 2.2897049348102883e-05
step: 270, loss: 0.00020711969409603626
step: 280, loss: 0.004669290967285633
step: 290, loss: 7.407142402371392e-05
step: 300, loss: 2.6488602088647895e-05
step: 310, loss: 6.522944204334635e-06
step: 320, loss: 5.192090611672029e-05
step: 330, loss: 0.0012344964779913425
step: 340, loss: 0.00081878702621907
step: 350, loss: 7.308968633878976e-06
step: 360, loss: 8.948037248046603e-06
step: 370, loss: 0.0005526047316379845
step: 380, loss: 0.0006971347611397505
step: 390, loss: 8.402786625083536e-05
step: 400, loss: 0.004320893436670303
step: 410, loss: 2.7670941562973894e-05
step: 420, loss: 1.2237187547725625e-05
step: 430, loss: 8.243995580414776e-06
step: 440, loss: 3.3976273698499426e-05
step: 450, loss: 7.536208158853697e-06
step: 460, loss: 0.010626988485455513
step: 470, loss: 8.694743883097544e-06
step: 480, loss: 6.91037212163792e-06
step: 490, loss: 7.599537639180198e-06
step: 500, loss: 0.0016802866011857986

```

step: 510, loss: 0.000705496349837631
step: 520, loss: 0.0008164438768289983
step: 530, loss: 0.0010670417686924338
/usr/local/lib/python3.7/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=0.9525598872710193, f1=0.947022972339428,
best_f1=0.954001839926403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.001845344551839e-06
step: 10, loss: 0.002109113149344921
step: 20, loss: 0.0017679572338238358
step: 30, loss: 0.0011326548410579562
step: 40, loss: 1.1995241948170587e-05
step: 50, loss: 0.00033224659273400903
step: 60, loss: 0.0013862629421055317
step: 70, loss: 1.1186804840690456e-05
step: 80, loss: 7.461691893695388e-06
step: 90, loss: 7.908720363047905e-06
step: 100, loss: 5.092450464871945e-06
step: 110, loss: 0.00016557784692849964
step: 120, loss: 7.163683676481014e-06
step: 130, loss: 1.5283912944141775e-05
step: 140, loss: 6.392555405909661e-06
step: 150, loss: 1.0657926395651884e-05
step: 160, loss: 5.923180651734583e-06
step: 170, loss: 2.6609292035573162e-05
step: 180, loss: 1.5239778804243542e-05
step: 190, loss: 0.0013399645686149597
step: 200, loss: 0.0007309236680157483
step: 210, loss: 1.6862937627593055e-05
step: 220, loss: 7.95714913692791e-06
step: 230, loss: 9.257265446649399e-06
step: 240, loss: 1.065411288436735e-05
step: 250, loss: 3.3261447242693976e-05
step: 260, loss: 8.251407962234225e-06
step: 270, loss: 7.696392458456103e-06
step: 280, loss: 7.1375834522768855e-06
step: 290, loss: 0.0010131036397069693
step: 300, loss: 6.765069883840624e-06
step: 310, loss: 0.002486619632691145
step: 320, loss: 7.174865459091961e-06
step: 330, loss: 0.00023471990425605327
step: 340, loss: 8.787886145000812e-06
step: 350, loss: 0.0011184990871697664
step: 360, loss: 0.00033916591200977564

```

```

step: 370, loss: 2.114215021720156e-05
step: 380, loss: 1.1387945960450452e-05
step: 390, loss: 1.681150206422899e-05
step: 400, loss: 0.0015349796740338206
step: 410, loss: 7.565997293568216e-06
step: 420, loss: 2.144686004612595e-05
step: 430, loss: 5.673581654264126e-06
step: 440, loss: 6.400011443474796e-06
step: 450, loss: 0.001916887704282999
step: 460, loss: 0.0007026362582109869
step: 470, loss: 6.8544877649401315e-06
step: 480, loss: 0.0024021996650844812
step: 490, loss: 6.651429430348799e-05
step: 500, loss: 0.001253367867320776
step: 510, loss: 6.582532023458043e-06
step: 520, loss: 2.2537675249623135e-05
step: 530, loss: 6.4000082602433395e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9534992954438704, f1=0.9480580252690688,
best_f1=0.954001839926403
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 [00:27, 208.31it/s]
load_f1 = 0.955637707948244
real_f1 = 0.9541454377026402
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:25, 171.90it/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: 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: 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: 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: 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: 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: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
```

```

packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: 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: 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: 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: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: 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=5b6bfa33c3e216d7810e580e25742abd48c44e7ec6a2bfc8c64bb27c279b9001
  Stored in directory: /tmp/pip-ephem-wheel-cache-
mz33kr3v/wheels/11/95/ba/2c36cc368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5

```


Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
Download and installation successful
You can now load the model via `spacy.load('en_core_web_lg')`

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

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

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

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

Downloading: 100% 501M/501M [00:07<00:00, 67.0MB/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: 0.4614316523075104
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.29545454545454547, f1=0.25316455696202533,
best_f1=0.25316455696202533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4841925799846649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3333333333333333, f1=0.28571428571428575,
best_f1=0.28571428571428575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4196210205554962
/usr/local/lib/python3.7/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=0.30434782608695654, f1=0.2857142857142857,
best_f1=0.28571428571428575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3357836902141571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5517241379310344, f1=0.55, best_f1=0.55
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2502564787864685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.7586206896551724, f1=0.6060606060606061,
best_f1=0.6060606060606061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1268058866262436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8571428571428571, f1=0.7142857142857143,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31151723861694336
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.7741935483870968,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1505563110113144
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.6923076923076924,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0709352195262909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8, f1=0.5454545454545454, best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28781262040138245
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.8275862068965518,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12085378915071487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8666666666666666, f1=0.7142857142857143,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016904689371585846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8666666666666666, f1=0.7999999999999999,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011594644747674465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9032258064516129, f1=0.8125000000000001,
best_f1=0.8125000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018098028376698494
/usr/local/lib/python3.7/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=0.8750000000000001, f1=0.8125000000000001,
best_f1=0.8125000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018990593031048775
/usr/local/lib/python3.7/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=0.8750000000000001, f1=0.8125000000000001,
best_f1=0.8125000000000001
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, 126301.01it/s]

/usr/local/lib/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 = 0.8148148148148148

real_f1 = 0.8571428571428571

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

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

4398it [00:30, 144.85it/s]

3.1.2 DBLP-ACM - Running the matcher

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

```

--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.5486199855804443
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4379050135612488
step: 20, loss: 0.5638028383255005
step: 30, loss: 0.28011828660964966
step: 40, loss: 0.32004544138908386
step: 50, loss: 0.5847017168998718
step: 60, loss: 0.17498408257961273
step: 70, loss: 0.17107851803302765
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 80, loss: 0.19831036031246185
step: 90, loss: 0.09304876625537872
step: 100, loss: 0.15195319056510925
step: 110, loss: 0.11054126173257828
step: 120, loss: 0.04825717210769653
step: 130, loss: 0.022747119888663292
step: 140, loss: 0.11145834624767303
step: 150, loss: 0.3215790092945099
step: 160, loss: 0.005533230025321245
step: 170, loss: 0.014871135354042053
step: 180, loss: 0.029402492567896843
step: 190, loss: 0.03340177237987518
step: 200, loss: 0.051542237401008606
step: 210, loss: 0.13507050275802612
step: 220, loss: 0.03600883111357689
step: 230, loss: 0.06422343850135803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9581497797356827, f1=0.9640449438202248,
best_f1=0.9640449438202248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007114645559340715
step: 10, loss: 0.041623689234256744
step: 20, loss: 0.013920962810516357
step: 30, loss: 0.002968662418425083
step: 40, loss: 0.021019408479332924
step: 50, loss: 0.001455108867958188
step: 60, loss: 0.002296861493960023
step: 70, loss: 0.01942306011915207
step: 80, loss: 0.015762396156787872
step: 90, loss: 0.009715475142002106
step: 100, loss: 0.027826771140098572
step: 110, loss: 0.015392077155411243
step: 120, loss: 0.00791031401604414
step: 130, loss: 0.020223628729581833
step: 140, loss: 0.07389072328805923
step: 150, loss: 0.0655662789940834
step: 160, loss: 0.029934007674455643
step: 170, loss: 0.011761181987822056
step: 180, loss: 0.0075485678389668465
step: 190, loss: 0.025617631152272224
step: 200, loss: 0.02228886830210686
step: 210, loss: 0.19178904592990875
step: 220, loss: 0.006476634182035923
step: 230, loss: 0.0016499991761520505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9752252252252253, f1=0.9715585893060296,
best_f1=0.9715585893060296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035693731158971786
step: 10, loss: 0.0010554929031059146
step: 20, loss: 0.020802441984415054
step: 30, loss: 0.006472621113061905
step: 40, loss: 0.027606049552559853
step: 50, loss: 0.012383121065795422
step: 60, loss: 0.0010325743351131678

```

```

step: 70, loss: 0.0006682200473733246
step: 80, loss: 0.0002820360823534429
step: 90, loss: 0.008438198827207088
step: 100, loss: 0.05131042003631592
step: 110, loss: 0.04826737940311432
step: 120, loss: 0.006789391860365868
step: 130, loss: 0.0742112472653389
step: 140, loss: 0.0005811208975501359
step: 150, loss: 0.04257219284772873
step: 160, loss: 0.0009515952551737428
step: 170, loss: 0.005331980995833874
step: 180, loss: 0.0010838761227205396
step: 190, loss: 0.04877743870019913
step: 200, loss: 0.04398578777909279
step: 210, loss: 0.006549449171870947
step: 220, loss: 0.0012147566303610802
step: 230, loss: 0.008772236295044422
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9809203142536477, f1=0.9809203142536477,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004123429302126169
step: 10, loss: 0.0015987810911610723
step: 20, loss: 0.0021157325245440006
step: 30, loss: 0.0005847717402502894
step: 40, loss: 0.011873300187289715
step: 50, loss: 0.058289844542741776
step: 60, loss: 0.006974917370826006
step: 70, loss: 0.019257228821516037
step: 80, loss: 0.0006333499914035201
step: 90, loss: 0.0013088133418932557
step: 100, loss: 0.0011200614972040057
step: 110, loss: 0.0013258361723273993
step: 120, loss: 0.11508958041667938
step: 130, loss: 0.001826587482355535
step: 140, loss: 0.0009541245526634157
step: 150, loss: 0.0016960144275799394
step: 160, loss: 0.0005347529659047723
step: 170, loss: 0.013206207193434238
step: 180, loss: 0.06501984596252441

```

```

step: 190, loss: 0.0009509860537946224
step: 200, loss: 0.012304456904530525
step: 210, loss: 0.0035930145531892776
step: 220, loss: 0.0002975761017296463
step: 230, loss: 0.04388988018035889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9810479375696767, f1=0.9820627802690582,
best_f1=0.9820627802690582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002884666435420513
step: 10, loss: 0.0044276961125433445
step: 20, loss: 0.10127949714660645
step: 30, loss: 0.002395081566646695
step: 40, loss: 0.0017663058824837208
step: 50, loss: 0.0006880878354422748
step: 60, loss: 0.0007924935780465603
step: 70, loss: 0.001327293342910707
step: 80, loss: 0.004649042151868343
step: 90, loss: 0.004200930241495371
step: 100, loss: 0.00037288962630555034
step: 110, loss: 0.0026481139939278364
step: 120, loss: 0.05914699658751488
step: 130, loss: 0.004172125365585089
step: 140, loss: 0.0006725971470586956
step: 150, loss: 0.0034267278388142586
step: 160, loss: 0.0003175045712850988
step: 170, loss: 0.0050877295434474945
step: 180, loss: 0.0013265195302665234
step: 190, loss: 0.005805641878396273
step: 200, loss: 0.0010958677157759666
step: 210, loss: 0.002229186473414302
step: 220, loss: 0.0012071311939507723
step: 230, loss: 0.0038100476376712322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9876819708846584, f1=0.9808773903262092,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010762031888589263
step: 10, loss: 0.00493033230304718
step: 20, loss: 0.0034238097723573446
step: 30, loss: 0.0010448605753481388
step: 40, loss: 0.0005313542787916958
step: 50, loss: 0.00044498094939626753
step: 60, loss: 0.02698398008942604
step: 70, loss: 0.013154859654605389
step: 80, loss: 0.003296618815511465
step: 90, loss: 0.012728327885270119
step: 100, loss: 0.0016372036188840866
step: 110, loss: 0.029165096580982208
step: 120, loss: 0.0021639359183609486
step: 130, loss: 0.004923987202346325
step: 140, loss: 0.0017505364958196878
step: 150, loss: 0.00820532999932766
step: 160, loss: 0.03518545255064964
step: 170, loss: 0.0011029258603230119
step: 180, loss: 0.0013960336800664663
step: 190, loss: 0.0010833840351551771
step: 200, loss: 0.004145604558289051
step: 210, loss: 0.0006735408678650856
step: 220, loss: 0.012010536156594753
step: 230, loss: 0.001313297776505351
/usr/local/lib/python3.7/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=0.984304932735426, f1=0.9819819819819819,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024908732157200575
step: 10, loss: 0.0009355123620480299
step: 20, loss: 0.0007847769884392619
step: 30, loss: 0.0016213557682931423
step: 40, loss: 0.00226286961697042
step: 50, loss: 0.0010195986833423376
step: 60, loss: 0.0012345591094344854
step: 70, loss: 0.0006756382645107806
step: 80, loss: 0.000695135910063982
step: 90, loss: 0.024783369153738022
step: 100, loss: 0.07317198812961578

```

```

step: 110, loss: 0.0031408420763909817
step: 120, loss: 0.0008009958546608686
step: 130, loss: 0.004083039239048958
step: 140, loss: 0.0001720525324344635
step: 150, loss: 0.007926827296614647
step: 160, loss: 0.00039955557440407574
step: 170, loss: 0.0019254212966188788
step: 180, loss: 0.0006386422901414335
step: 190, loss: 0.00101001956500113
step: 200, loss: 0.0067207589745521545
step: 210, loss: 0.20378194749355316
step: 220, loss: 0.0018700516084209085
step: 230, loss: 0.005335867404937744
/usr/local/lib/python3.7/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=0.9852104664391355, f1=0.977116704805492,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005114930681884289
step: 10, loss: 0.008537685498595238
step: 20, loss: 0.002347043016925454
step: 30, loss: 0.0017840116051957011
step: 40, loss: 0.0022208362352102995
step: 50, loss: 0.006706735584884882
step: 60, loss: 0.0008110355702228844
step: 70, loss: 0.00043892176472581923
step: 80, loss: 0.0003884669567923993
step: 90, loss: 0.02945110946893692
step: 100, loss: 0.0004261407011654228
step: 110, loss: 0.003343351650983095
step: 120, loss: 0.000379729550331831
step: 130, loss: 0.016893625259399414
step: 140, loss: 0.0004212422645650804
step: 150, loss: 0.12189753353595734
step: 160, loss: 0.0005191604141145945
step: 170, loss: 0.028549835085868835
step: 180, loss: 0.00043541949708014727
step: 190, loss: 0.004033468663692474
step: 200, loss: 0.004449853673577309
step: 210, loss: 0.02673056721687317
step: 220, loss: 0.0021082167513668537
step: 230, loss: 0.001020737225189805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898074745186863, f1=0.9864253393665158,
best_f1=0.9864253393665158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005304568912833929
step: 10, loss: 0.00310192396864295
step: 20, loss: 0.0010616440558806062
step: 30, loss: 0.0008653425611555576
step: 40, loss: 0.0007109635043889284
step: 50, loss: 0.002130882814526558
step: 60, loss: 0.0007586394203826785
step: 70, loss: 0.07874690741300583
step: 80, loss: 0.00025509498664177954
step: 90, loss: 0.08734161406755447
step: 100, loss: 0.007981612347066402
step: 110, loss: 0.00027999337180517614
step: 120, loss: 0.04198422655463219
step: 130, loss: 0.00048716345918364823
step: 140, loss: 0.0005204659537412226
step: 150, loss: 0.004538453184068203
step: 160, loss: 0.0007722312584519386
step: 170, loss: 0.0002867799485102296
step: 180, loss: 0.0015933022368699312
step: 190, loss: 0.0002348864945815876
step: 200, loss: 0.0003268294676672667
step: 210, loss: 0.0009920424781739712
step: 220, loss: 0.00039500428829342127
step: 230, loss: 0.0003223066742066294
/usr/local/lib/python3.7/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=0.9887133182844244, f1=0.9852774631936579,
best_f1=0.9864253393665158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004518468340393156
step: 10, loss: 0.0007750988588668406
step: 20, loss: 0.00039909896440804005
step: 30, loss: 0.00028244819259271026
step: 40, loss: 0.001337121007964015
step: 50, loss: 0.0007385819335468113
step: 60, loss: 0.004566386342048645

```

```

step: 70, loss: 0.0035776845179498196
step: 80, loss: 0.0003516036958899349
step: 90, loss: 0.00026056531351059675
step: 100, loss: 0.0002719162730500102
step: 110, loss: 0.0006423068698495626
step: 120, loss: 0.00029931747121736407
step: 130, loss: 0.0004332715179771185
step: 140, loss: 0.00027557360590435565
step: 150, loss: 0.0005306251696310937
step: 160, loss: 0.00017746428784448653
step: 170, loss: 0.00035345120704732835
step: 180, loss: 0.0006814555381424725
step: 190, loss: 0.00042249064426869154
step: 200, loss: 0.000688976957462728
step: 210, loss: 0.0004824287025257945
step: 220, loss: 0.007654554210603237
step: 230, loss: 0.00027195236179977655
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898534385569334, f1=0.9864864864864865,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019075747695751488
step: 10, loss: 0.00044288806384429336
step: 20, loss: 0.0003718246298376471
step: 30, loss: 0.0004358105652499944
step: 40, loss: 0.0007440616027452052
step: 50, loss: 0.00035191542701795697
step: 60, loss: 0.019040580838918686
step: 70, loss: 0.009934410452842712
step: 80, loss: 0.00026719519519247115
step: 90, loss: 0.18992963433265686
step: 100, loss: 0.0023874747566878796
step: 110, loss: 0.00097014190396294
step: 120, loss: 0.0007563233375549316
step: 130, loss: 0.0005548063199967146
step: 140, loss: 0.0018708814168348908
step: 150, loss: 0.0006585908122360706
step: 160, loss: 0.006727053318172693
step: 170, loss: 0.0028557872865349054
step: 180, loss: 0.0013758688000962138

```

```

step: 190, loss: 0.0004951037699356675
step: 200, loss: 0.006208289414644241
step: 210, loss: 0.0013242131099104881
step: 220, loss: 0.005393852014094591
step: 230, loss: 0.0003092741244472563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9876543209876544, f1=0.9820224719101124,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004860307672061026
step: 10, loss: 0.00027960806619375944
step: 20, loss: 0.01110239140689373
step: 30, loss: 0.028035828843712807
step: 40, loss: 0.0008895836654119194
step: 50, loss: 0.004181985277682543
step: 60, loss: 0.0006642970838584006
step: 70, loss: 0.0005801856750622392
step: 80, loss: 0.00015724792319815606
step: 90, loss: 0.002079753205180168
step: 100, loss: 0.00024086012854240835
step: 110, loss: 0.00016168606816790998
step: 120, loss: 0.002297869883477688
step: 130, loss: 0.0005445068818517029
step: 140, loss: 0.0008747664978727698
step: 150, loss: 0.00034440140007063746
step: 160, loss: 0.0004343477776274085
step: 170, loss: 0.0020990464836359024
step: 180, loss: 0.00033674348378553987
step: 190, loss: 0.0038387873210012913
step: 200, loss: 0.000280263222521171
step: 210, loss: 0.0004719517019111663
step: 220, loss: 0.013761937618255615
step: 230, loss: 0.0004055506724398583
/usr/local/lib/python3.7/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=0.9875706214689265,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011024451814591885
step: 10, loss: 0.00023307382070925087
step: 20, loss: 0.001554465270601213

```



```

step: 30, loss: 0.00031155371107161045
step: 40, loss: 0.0013811031822115183
step: 50, loss: 0.003976375330239534
step: 60, loss: 0.00038428764673881233
step: 70, loss: 0.0013868095120415092
step: 80, loss: 0.0013007005909457803
step: 90, loss: 0.00044551517930813134
step: 100, loss: 0.001742169144563377
step: 110, loss: 0.0012951522367075086
step: 120, loss: 0.0007940015057101846
step: 130, loss: 0.0002557167608756572
step: 140, loss: 0.0003720156382769346
step: 150, loss: 7.863094651838765e-05
step: 160, loss: 0.0012081917375326157
step: 170, loss: 0.0027467021718621254
step: 180, loss: 0.03714153915643692
step: 190, loss: 0.0003365858574397862
step: 200, loss: 6.197447510203347e-05
step: 210, loss: 0.002510341349989176
step: 220, loss: 0.0003319628303870559
step: 230, loss: 0.0003017802373506129
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9831271091113611,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011720426118699834
step: 10, loss: 9.700866212369874e-05
step: 20, loss: 0.00018062637536786497
step: 30, loss: 0.000420232187025249
step: 40, loss: 9.596822928870097e-05
step: 50, loss: 0.00021942002058494836
step: 60, loss: 0.00025822402676567435
step: 70, loss: 0.0001701847795629874
step: 80, loss: 0.0005922128912061453
step: 90, loss: 0.0026235547848045826
step: 100, loss: 0.0012331211473792791
step: 110, loss: 0.0002669971436262131
step: 120, loss: 4.992468166165054e-05
step: 130, loss: 0.00021439902775455266
step: 140, loss: 0.0002227427321486175
step: 150, loss: 4.923560481984168e-05
step: 160, loss: 0.0004889875999651849
step: 170, loss: 0.00024461362045258284
step: 180, loss: 0.00014828750863671303

```

```

step: 190, loss: 0.0001804341736715287
step: 200, loss: 0.0009600965422578156
step: 210, loss: 0.00015999893366824836
step: 220, loss: 0.00018485941109247506
step: 230, loss: 9.395687811775133e-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=0.9876543209876544, f1=0.9842696629213483,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0176670141518116
step: 10, loss: 0.0002765968383755535
step: 20, loss: 0.0021470363717526197
step: 30, loss: 0.0001607035956112668
step: 40, loss: 0.000250336219323799
step: 50, loss: 0.00013635223149321973
step: 60, loss: 0.058985378593206406
step: 70, loss: 0.0003145758237224072
step: 80, loss: 0.00013582811516243964
step: 90, loss: 0.00020932205370627344
step: 100, loss: 9.981414768844843e-05
step: 110, loss: 0.0001835139119066298
step: 120, loss: 0.02850107103586197
step: 130, loss: 0.00014108780305832624
step: 140, loss: 0.0070040347054600716
step: 150, loss: 0.00017452608153689653
step: 160, loss: 0.010714912787079811
step: 170, loss: 6.284717528615147e-05
step: 180, loss: 0.002314545912668109
step: 190, loss: 0.00032210751669481397
step: 200, loss: 0.0014550368068739772
step: 210, loss: 0.011503443121910095
step: 220, loss: 0.00015882670413702726
step: 230, loss: 0.00017558957915753126
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9842696629213483,
best_f1=0.9864864864864865
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 [00:17, 144.60it/s]
load_f1 = 0.9898762654668166
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 [00:33, 130.70it/s]

```

3.1.3 DBLP-GoogleScholar - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.6188238859176636

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.45748141407966614
step: 20, loss: 0.3040255904197693
step: 30, loss: 0.3328167498111725
step: 40, loss: 0.33275845646858215
step: 50, loss: 0.3507930636405945
step: 60, loss: 0.09721757471561432
step: 70, loss: 0.3080408573150635
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.10292492061853409
step: 90, loss: 0.213617205619812
step: 100, loss: 0.1406605839729309
step: 110, loss: 0.09398798644542694
step: 120, loss: 0.14416372776031494
step: 130, loss: 0.11904973536729813
step: 140, loss: 0.15404047071933746
step: 150, loss: 0.08746249973773956
step: 160, loss: 0.17654626071453094
step: 170, loss: 0.04822002351284027
step: 180, loss: 0.07108522951602936
step: 190, loss: 0.06202555075287819
step: 200, loss: 0.022978434339165688
step: 210, loss: 0.03428840637207031
step: 220, loss: 0.18238967657089233
step: 230, loss: 0.1626981645822525
step: 240, loss: 0.04193751513957977
step: 250, loss: 0.07087667286396027
step: 260, loss: 0.23446007072925568
step: 270, loss: 0.3180544078350067
step: 280, loss: 0.06264682859182358
step: 290, loss: 0.13361482322216034
step: 300, loss: 0.02772160992026329
step: 310, loss: 0.13553541898727417
step: 320, loss: 0.12579916417598724
step: 330, loss: 0.16848525404930115
step: 340, loss: 0.5909309387207031
step: 350, loss: 0.12115619331598282
step: 360, loss: 0.06721813231706619
step: 370, loss: 0.07620566338300705
step: 380, loss: 0.1867814064025879
step: 390, loss: 0.021803978830575943
step: 400, loss: 0.053443364799022675
step: 410, loss: 0.25571051239967346
step: 420, loss: 0.011972647160291672
step: 430, loss: 0.016778474673628807
step: 440, loss: 0.05709170922636986
step: 450, loss: 0.09803558140993118

```

step: 460, loss: 0.021734731271862984
step: 470, loss: 0.03950239717960358
step: 480, loss: 0.16749480366706848
step: 490, loss: 0.17489252984523773
step: 500, loss: 0.02054774947464466
step: 510, loss: 0.06687886267900467
step: 520, loss: 0.11958061903715134
step: 530, loss: 0.007810610346496105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9409576940957695, f1=0.9424326833797586,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033059101551771164
step: 10, loss: 0.17456857860088348
step: 20, loss: 0.054417677223682404
step: 30, loss: 0.05688678100705147
step: 40, loss: 0.07831283658742905
step: 50, loss: 0.14657726883888245
step: 60, loss: 0.05059288069605827
step: 70, loss: 0.029589150100946426
step: 80, loss: 0.027354411780834198
step: 90, loss: 0.003969786688685417
step: 100, loss: 0.10496693104505539
step: 110, loss: 0.01528165489435196
step: 120, loss: 0.16002368927001953
step: 130, loss: 0.04230896756052971
step: 140, loss: 0.03785058483481407
step: 150, loss: 0.013276764191687107
step: 160, loss: 0.03784399852156639
step: 170, loss: 0.05956389755010605
step: 180, loss: 0.035180315375328064
step: 190, loss: 0.06114790216088295
step: 200, loss: 0.19222237169742584
step: 210, loss: 0.06273642182350159
step: 220, loss: 0.002632268937304616
step: 230, loss: 0.016064785420894623
step: 240, loss: 0.1399652659893036
step: 250, loss: 0.0374675951898098
step: 260, loss: 0.017415577545762062
step: 270, loss: 0.01779817044734955

```

```

step: 280, loss: 0.05948847532272339
step: 290, loss: 0.052440956234931946
step: 300, loss: 0.024114197120070457
step: 310, loss: 0.039753515273332596
step: 320, loss: 0.049869731068611145
step: 330, loss: 0.00801790226250887
step: 340, loss: 0.03901410847902298
step: 350, loss: 0.0049805608578026295
step: 360, loss: 0.05084194242954254
step: 370, loss: 0.0010369475930929184
step: 380, loss: 0.2754332423210144
step: 390, loss: 0.014988232403993607
step: 400, loss: 0.02658967860043049
step: 410, loss: 0.12238039821386337
step: 420, loss: 0.05011267587542534
step: 430, loss: 0.2315954864025116
step: 440, loss: 0.009157788008451462
step: 450, loss: 0.061011169105768204
step: 460, loss: 0.03923499584197998
step: 470, loss: 0.026957562193274498
step: 480, loss: 0.00202798075042665
step: 490, loss: 0.06501718610525131
step: 500, loss: 0.008139543235301971
step: 510, loss: 0.009283533319830894
step: 520, loss: 0.3753072917461395
step: 530, loss: 0.019952522590756416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9507593189139439, f1=0.9442139234670355,
best_f1=0.9442139234670355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07189963012933731
step: 10, loss: 0.04243973270058632
step: 20, loss: 0.00730513921007514
step: 30, loss: 0.13067932426929474
step: 40, loss: 0.06018224358558655
step: 50, loss: 0.006223083008080721
step: 60, loss: 0.010922790504992008
step: 70, loss: 0.006947140675038099
step: 80, loss: 0.0017147164326161146
step: 90, loss: 0.0041458806954324245

```

```
step: 100, loss: 0.03271936625242233
step: 110, loss: 0.03179367631673813
step: 120, loss: 0.13798177242279053
step: 130, loss: 0.04851031303405762
step: 140, loss: 0.007455938030034304
step: 150, loss: 0.010710753500461578
step: 160, loss: 0.005040369462221861
step: 170, loss: 0.002562546404078603
step: 180, loss: 0.03645377606153488
step: 190, loss: 0.004483216442167759
step: 200, loss: 0.03842606395483017
step: 210, loss: 0.04905683919787407
step: 220, loss: 0.026550976559519768
step: 230, loss: 0.03160906210541725
step: 240, loss: 0.06851746886968613
step: 250, loss: 0.08844010531902313
step: 260, loss: 0.09310203045606613
step: 270, loss: 0.0035310874227434397
step: 280, loss: 0.0009889926295727491
step: 290, loss: 0.0062406109645962715
step: 300, loss: 0.06403665989637375
step: 310, loss: 0.0674070417881012
step: 320, loss: 0.19015629589557648
step: 330, loss: 0.00406182836741209
step: 340, loss: 0.0062524303793907166
step: 350, loss: 0.12119261920452118
step: 360, loss: 0.010217337869107723
step: 370, loss: 0.0209105983376503
step: 380, loss: 0.007122586946934462
step: 390, loss: 0.005853199865669012
step: 400, loss: 0.09227554500102997
step: 410, loss: 0.01566622406244278
step: 420, loss: 0.024919239804148674
step: 430, loss: 0.006246154196560383
step: 440, loss: 0.18947918713092804
step: 450, loss: 0.15538521111011505
step: 460, loss: 0.034029316157102585
step: 470, loss: 0.1295909434556961
step: 480, loss: 0.060047004371881485
step: 490, loss: 0.025902867317199707
step: 500, loss: 0.02155821956694126
step: 510, loss: 0.02491661347448826
step: 520, loss: 0.012842914089560509
step: 530, loss: 0.01292046532034874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```


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=0.9549632352941175, f1=0.9471264367816092,

best_f1=0.9471264367816092

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

An input tensor was not cuda.

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

step: 0, loss: 0.02345062606036663

step: 10, loss: 0.005692279897630215

step: 20, loss: 0.060018137097358704

step: 30, loss: 0.0103401318192482

step: 40, loss: 0.0370112806558609

step: 50, loss: 0.10564721375703812

step: 60, loss: 0.008157031610608101

step: 70, loss: 0.002214468317106366

step: 80, loss: 0.0018171854317188263

step: 90, loss: 0.043084483593702316

step: 100, loss: 0.0030698482878506184

step: 110, loss: 0.00461615389212966

step: 120, loss: 0.002855336293578148

step: 130, loss: 0.1411115974187851

step: 140, loss: 0.03690319135785103

step: 150, loss: 0.027547761797904968

step: 160, loss: 0.011743666604161263

step: 170, loss: 0.013012785464525223

step: 180, loss: 0.17111563682556152

step: 190, loss: 0.025553347542881966

step: 200, loss: 0.013533015735447407

step: 210, loss: 0.0038756721187382936

step: 220, loss: 0.0003411217185202986

step: 230, loss: 0.029655465856194496

step: 240, loss: 0.007787054870277643

step: 250, loss: 0.12703044712543488

step: 260, loss: 0.006427394226193428

step: 270, loss: 0.03510052338242531

step: 280, loss: 0.0054396032355725765

step: 290, loss: 0.055475618690252304

step: 300, loss: 0.0012506042839959264

step: 310, loss: 0.0021630297414958477

step: 320, loss: 0.16588819026947021

step: 330, loss: 0.006201880984008312

step: 340, loss: 0.009691094979643822

step: 350, loss: 0.09562762081623077

step: 360, loss: 0.02844015136361122

step: 370, loss: 0.007719664368778467

step: 380, loss: 0.009729469195008278

step: 390, loss: 0.0005497772945091128

```

step: 400, loss: 0.008740723133087158
step: 410, loss: 0.0012100358726456761
step: 420, loss: 0.010457322932779789
step: 430, loss: 0.015325370244681835
step: 440, loss: 0.06431291252374649
step: 450, loss: 0.02455870434641838
step: 460, loss: 0.04026792570948601
step: 470, loss: 0.0006724843988195062
step: 480, loss: 0.041558846831321716
step: 490, loss: 0.004745832644402981
step: 500, loss: 0.04204915463924408
step: 510, loss: 0.05200415849685669
step: 520, loss: 0.002886503469198942
step: 530, loss: 0.10207132995128632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9563994374120957, f1=0.9477154969382948,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010380018502473831
step: 10, loss: 0.02398526296019554
step: 20, loss: 0.0029220760334283113
step: 30, loss: 0.054797977209091187
step: 40, loss: 0.0013292721705511212
step: 50, loss: 0.13931629061698914
step: 60, loss: 0.048388246446847916
step: 70, loss: 0.001942927367053926
step: 80, loss: 0.00042368806316517293
step: 90, loss: 0.051640138030052185
step: 100, loss: 0.06672630459070206
step: 110, loss: 0.017427289858460426
step: 120, loss: 0.054812829941511154
step: 130, loss: 0.007838202640414238
step: 140, loss: 0.02014477178454399
step: 150, loss: 0.044347744435071945
step: 160, loss: 0.017334651201963425
step: 170, loss: 0.09315066039562225
step: 180, loss: 0.015187532640993595
step: 190, loss: 0.003763003973290324
step: 200, loss: 0.015732398256659508
step: 210, loss: 0.03903349116444588

```

```
step: 220, loss: 0.000925532600376755
step: 230, loss: 0.0005532811628654599
step: 240, loss: 0.015244237147271633
step: 250, loss: 0.1014786809682846
step: 260, loss: 0.0012565483339130878
step: 270, loss: 0.02631022222340107
step: 280, loss: 0.06597919762134552
step: 290, loss: 0.013319611549377441
step: 300, loss: 0.012093092314898968
step: 310, loss: 0.014143078587949276
step: 320, loss: 0.11228140443563461
step: 330, loss: 0.0018806520383805037
step: 340, loss: 0.0508219450712204
step: 350, loss: 0.0014481086982414126
step: 360, loss: 0.0007891635177657008
step: 370, loss: 0.009198804385960102
step: 380, loss: 0.0017277891747653484
step: 390, loss: 0.004841525107622147
step: 400, loss: 0.0007683911244384944
step: 410, loss: 0.02110128290951252
step: 420, loss: 0.21311385929584503
step: 430, loss: 0.003793608397245407
step: 440, loss: 0.0014735382283106446
step: 450, loss: 0.01623423583805561
step: 460, loss: 0.05506276339292526
step: 470, loss: 0.0343208983540535
step: 480, loss: 0.004263011272996664
step: 490, loss: 0.004151578526943922
step: 500, loss: 0.03644121438264847
step: 510, loss: 0.0022269829642027617
step: 520, loss: 0.09822910279035568
step: 530, loss: 0.013785106129944324
/usr/local/lib/python3.7/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=0.9520260829063809, f1=0.9505135387488329,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06449758261442184
step: 10, loss: 0.0002565899631008506
step: 20, loss: 0.00034448117366991937
step: 30, loss: 0.0010343100875616074
step: 40, loss: 0.0004939345526508987
step: 50, loss: 0.003402572125196457
step: 60, loss: 0.019526753574609756
step: 70, loss: 0.0010615199571475387
```

```
step: 80, loss: 0.0002999919524881989
step: 90, loss: 0.00031363009475171566
step: 100, loss: 0.2830575108528137
step: 110, loss: 0.0018841004930436611
step: 120, loss: 0.01083350833505392
step: 130, loss: 0.0006775836227461696
step: 140, loss: 0.007083290722221136
step: 150, loss: 0.00506451353430748
step: 160, loss: 0.15477026998996735
step: 170, loss: 0.00559987174347043
step: 180, loss: 0.005231175571680069
step: 190, loss: 0.027040984481573105
step: 200, loss: 0.015718627721071243
step: 210, loss: 0.0006985141080804169
step: 220, loss: 0.0012459808494895697
step: 230, loss: 0.024120589718222618
step: 240, loss: 0.200161412358284
step: 250, loss: 0.05861228331923485
step: 260, loss: 0.00765560707077384
step: 270, loss: 0.007972870022058487
step: 280, loss: 0.0042240675538778305
step: 290, loss: 0.0003857699630316347
step: 300, loss: 0.0027206146623939276
step: 310, loss: 0.06508932262659073
step: 320, loss: 0.00023344634973909706
step: 330, loss: 0.0061560943722724915
step: 340, loss: 0.0006731277680955827
step: 350, loss: 0.0102081298828125
step: 360, loss: 0.031234135851264
step: 370, loss: 0.01470436342060566
step: 380, loss: 0.001047411235049367
step: 390, loss: 0.002805189462378621
step: 400, loss: 0.0023612407967448235
step: 410, loss: 0.004474494140595198
step: 420, loss: 0.016530033200979233
step: 430, loss: 0.003555819159373641
step: 440, loss: 0.07105087488889694
step: 450, loss: 0.2142612189054489
step: 460, loss: 0.010401374660432339
step: 470, loss: 0.0006703269318677485
step: 480, loss: 0.0024548026267439127
step: 490, loss: 0.004490723833441734
step: 500, loss: 0.0036146328784525394
step: 510, loss: 0.02795296348631382
step: 520, loss: 0.004556834697723389
step: 530, loss: 0.015245635993778706
/usr/local/lib/python3.7/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=0.9473193473193473, f1=0.948162111215834,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008070646785199642
step: 10, loss: 0.0029796359594911337
step: 20, loss: 0.009754146449267864
step: 30, loss: 0.011287808418273926
step: 40, loss: 0.004972292575985193
step: 50, loss: 0.002676036674529314
step: 60, loss: 0.004410037770867348
step: 70, loss: 0.0005269805551506579
step: 80, loss: 0.00048718886682763696
step: 90, loss: 0.00023537129163742065
step: 100, loss: 0.0006920526502653956
step: 110, loss: 0.00013712586951442063
step: 120, loss: 0.00029671602533198893
step: 130, loss: 0.0003290344320703298
step: 140, loss: 0.00030165398493409157
step: 150, loss: 0.0020676793064922094
step: 160, loss: 0.0003404525632504374
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.007817627862095833
step: 180, loss: 0.12247568368911743
step: 190, loss: 0.013522186316549778
step: 200, loss: 0.0009051637607626617
step: 210, loss: 0.0007638518000021577
step: 220, loss: 0.00011251930118305609
step: 230, loss: 0.0006780902622267604
step: 240, loss: 0.03138977289199829
step: 250, loss: 0.07515682280063629
step: 260, loss: 0.0019447029335424304
step: 270, loss: 0.00048791273729875684
step: 280, loss: 0.004603539127856493
step: 290, loss: 0.00019849147065542638
step: 300, loss: 0.0001199152902700007
step: 310, loss: 0.0042349048890173435
step: 320, loss: 0.09647223353385925
step: 330, loss: 0.00023675763804931194
step: 340, loss: 0.001238823402673006
step: 350, loss: 0.0006439696298912168
step: 360, loss: 0.003867338178679347
step: 370, loss: 0.01015226636081934
step: 380, loss: 0.0009061670280061662
step: 390, loss: 0.006730799563229084
step: 400, loss: 0.0017704892670735717

```

```
step: 410, loss: 0.00015457463450729847
step: 420, loss: 0.06306565552949905
step: 430, loss: 0.00016010957187972963
step: 440, loss: 0.0012321901740506291
step: 450, loss: 0.00028740064590238035
step: 460, loss: 0.0001949865254573524
step: 470, loss: 0.1706565022468567
step: 480, loss: 0.009257360361516476
step: 490, loss: 0.00021235315944068134
step: 500, loss: 0.0015529809752479196
step: 510, loss: 0.0005912821507081389
step: 520, loss: 0.0011286578373983502
step: 530, loss: 0.00040313063072972
/usr/local/lib/python3.7/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=0.9435370975268316, f1=0.939679547596607,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008437586948275566
step: 10, loss: 0.020124971866607666
step: 20, loss: 0.0005280346958898008
step: 30, loss: 0.00021927498164586723
step: 40, loss: 0.0003085437056142837
step: 50, loss: 0.00034391070948913693
step: 60, loss: 0.00031891666003502905
step: 70, loss: 0.0006460344884544611
step: 80, loss: 0.0015380613040179014
step: 90, loss: 0.0005343104712665081
step: 100, loss: 0.00589198200032115
step: 110, loss: 0.0003160380001645535
step: 120, loss: 0.001767642330378294
step: 130, loss: 0.01313470397144556
step: 140, loss: 0.00028173584723845124
step: 150, loss: 0.0571293942630291
step: 160, loss: 0.06328719109296799
step: 170, loss: 0.05656825006008148
step: 180, loss: 0.02682131715118885
step: 190, loss: 0.025991789996623993
step: 200, loss: 0.0033382843248546124
step: 210, loss: 0.0670192763209343
step: 220, loss: 0.00015577803424093872
step: 230, loss: 0.045480914413928986
step: 240, loss: 0.022954629734158516
step: 250, loss: 0.0005414113402366638
step: 260, loss: 0.0011736316373571754
```

```

step: 270, loss: 0.0004545266565401107
step: 280, loss: 0.0016076307510957122
step: 290, loss: 0.0009620212367735803
step: 300, loss: 8.468722808174789e-05
step: 310, loss: 0.0007176075014285743
step: 320, loss: 0.0008919679094105959
step: 330, loss: 0.0004064972454216331
step: 340, loss: 0.17547692358493805
step: 350, loss: 0.00039885807200334966
step: 360, loss: 0.04921136051416397
step: 370, loss: 0.007586474530398846
step: 380, loss: 0.000386832223739475
step: 390, loss: 0.013486308045685291
step: 400, loss: 0.002161305397748947
step: 410, loss: 0.004333114717155695
step: 420, loss: 0.001308559556491673
step: 430, loss: 0.0031588515266776085
step: 440, loss: 0.0012820437550544739
step: 450, loss: 0.0004804680065717548
step: 460, loss: 0.01131447684019804
step: 470, loss: 0.05943167582154274
step: 480, loss: 0.13640868663787842
step: 490, loss: 0.004261863883584738
step: 500, loss: 0.0012658372288569808
step: 510, loss: 0.002355037024244666
step: 520, loss: 0.0009746336727403104
step: 530, loss: 0.00033603451447561383
/usr/local/lib/python3.7/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=0.9477093937991671, f1=0.9471221338324755,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021047668997198343
step: 10, loss: 0.01636907458305359
step: 20, loss: 0.004620634485036135
step: 30, loss: 0.05008329823613167
step: 40, loss: 0.00027787324506789446
step: 50, loss: 0.0016003185883164406
step: 60, loss: 0.0010111292358487844
step: 70, loss: 0.16467006504535675
step: 80, loss: 0.015855703502893448
step: 90, loss: 0.06405754387378693
step: 100, loss: 0.0010241646086797118
step: 110, loss: 0.008951550349593163
step: 120, loss: 0.00024088460486382246

```

```

step: 130, loss: 0.00024775968631729484
step: 140, loss: 0.005431056953966618
step: 150, loss: 0.0001399677712470293
step: 160, loss: 0.0001741543528623879
step: 170, loss: 0.005344468168914318
step: 180, loss: 0.00011775989696616307
step: 190, loss: 7.666827877983451e-05
step: 200, loss: 0.00014068535529077053
step: 210, loss: 7.094401371432468e-05
step: 220, loss: 0.0018172778654843569
step: 230, loss: 0.001145757851190865
step: 240, loss: 7.611604814883322e-05
step: 250, loss: 0.0007545856642536819
step: 260, loss: 0.0015253758756443858
step: 270, loss: 0.03462657332420349
step: 280, loss: 0.017108792439103127
step: 290, loss: 0.00022297730902209878
step: 300, loss: 0.0010304481256753206
step: 310, loss: 0.02480117417871952
step: 320, loss: 0.03599517047405243
step: 330, loss: 0.0005009883316233754
step: 340, loss: 0.0013541245134547353
step: 350, loss: 0.022439874708652496
step: 360, loss: 0.0001855390873970464
step: 370, loss: 0.019887734204530716
step: 380, loss: 0.023680374026298523
step: 390, loss: 0.00028919108444824815
step: 400, loss: 0.0008365868707187474
step: 410, loss: 0.0034764311276376247
step: 420, loss: 0.00016383503680117428
step: 430, loss: 0.005185004323720932
step: 440, loss: 0.031185034662485123
step: 450, loss: 0.046942081302404404
step: 460, loss: 0.0001352698600385338
step: 470, loss: 0.0028069124091416597
step: 480, loss: 0.00011694486602209508
step: 490, loss: 0.0004014337609987706
step: 500, loss: 0.0027923877350986004
step: 510, loss: 0.0012577230809256434
step: 520, loss: 0.088225357234478
step: 530, loss: 0.0747952088713646
/usr/local/lib/python3.7/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=0.9496470588235294, f1=0.9479659413434247,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007014806033112109
step: 10, loss: 0.002577441046014428
step: 20, loss: 0.00040976126911118627
step: 30, loss: 0.0011690394021570683
step: 40, loss: 0.0003812774666585028
step: 50, loss: 0.0006909450166858733
step: 60, loss: 0.00911769736558199
step: 70, loss: 0.00010493527224753052
step: 80, loss: 7.551432645414025e-05
step: 90, loss: 7.174295751610771e-05
step: 100, loss: 0.003206136403605342
step: 110, loss: 0.013619635254144669
step: 120, loss: 0.0011409753933548927
step: 130, loss: 0.0016711876960471272
step: 140, loss: 0.00011767403339035809
step: 150, loss: 0.0008023447589948773
step: 160, loss: 0.03826112672686577
step: 170, loss: 3.542984268278815e-05
step: 180, loss: 0.00016866724763531238
step: 190, loss: 3.149216354358941e-05
step: 200, loss: 8.136266842484474e-05
step: 210, loss: 0.010190131142735481
step: 220, loss: 8.395307668251917e-05
step: 230, loss: 4.098728459211998e-05
step: 240, loss: 5.79413135710638e-05
step: 250, loss: 0.0017425876576453447
step: 260, loss: 0.0007276409887708724
step: 270, loss: 7.572754111606628e-05
step: 280, loss: 0.015420781448483467
step: 290, loss: 0.00011207479838049039
step: 300, loss: 0.0004883162910118699
step: 310, loss: 0.07383985072374344
step: 320, loss: 0.0011426148703321815
step: 330, loss: 0.004420161712914705
step: 340, loss: 0.0005616036360152066
step: 350, loss: 0.002624941524118185
step: 360, loss: 3.946623473893851e-05
step: 370, loss: 0.00022589584114030004
step: 380, loss: 0.0051671895198524
step: 390, loss: 2.3189300918602385e-05
step: 400, loss: 4.60761075373739e-05
step: 410, loss: 0.006458848714828491
step: 420, loss: 2.8031621695845388e-05
step: 430, loss: 2.6071100364788435e-05
step: 440, loss: 0.002945445943623781
step: 450, loss: 0.00035425741225481033
step: 460, loss: 2.96298494504299e-05
```

```

step: 470, loss: 0.00046683649998158216
step: 480, loss: 0.00010980905062751845
step: 490, loss: 4.872853241977282e-05
step: 500, loss: 0.008138788864016533
step: 510, loss: 2.1442159777507186e-05
step: 520, loss: 0.0006214614841155708
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 530, loss: 0.00038371802656911314
/usr/local/lib/python3.7/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=0.942008486562942, f1=0.9374999999999999,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9816015739925206e-05
step: 10, loss: 0.3764384090900421
step: 20, loss: 0.00017152755754068494
step: 30, loss: 0.0004777610010933131
step: 40, loss: 0.015898585319519043
step: 50, loss: 0.0001435368467355147
step: 60, loss: 6.0286674852250144e-05
step: 70, loss: 9.210805001202971e-05
step: 80, loss: 4.474955494515598e-05
step: 90, loss: 0.0008260383037850261
step: 100, loss: 0.00012061007873853669
step: 110, loss: 3.842134901788086e-05
step: 120, loss: 8.503460412612185e-05
step: 130, loss: 4.120293306186795e-05
step: 140, loss: 4.398930832394399e-05
step: 150, loss: 0.00047737525892443955
step: 160, loss: 0.00026376370806246996
step: 170, loss: 6.29658970865421e-05
step: 180, loss: 4.379880920168944e-05
step: 190, loss: 0.0012511418899521232
step: 200, loss: 5.220578168518841e-05
step: 210, loss: 0.0002128887572325766
step: 220, loss: 0.025775039568543434
step: 230, loss: 9.158676402876154e-05
step: 240, loss: 0.005341660231351852
step: 250, loss: 5.963830335531384e-05
step: 260, loss: 0.0055381907150149345
step: 270, loss: 6.286869756877422e-05
step: 280, loss: 0.00012285906996112317
step: 290, loss: 6.81313467794098e-05
step: 300, loss: 0.00041704505565576255
step: 310, loss: 0.15517404675483704

```

```

step: 320, loss: 7.846687367418781e-05
step: 330, loss: 0.0009084987686946988
step: 340, loss: 0.0010820202296599746
step: 350, loss: 0.00013436190783977509
step: 360, loss: 0.0028909833636134863
step: 370, loss: 0.0015088807558640838
step: 380, loss: 0.0070296465419232845
step: 390, loss: 0.011647133156657219
step: 400, loss: 0.005936469417065382
step: 410, loss: 0.002117990516126156
step: 420, loss: 0.04392467811703682
step: 430, loss: 0.002883083885535598
step: 440, loss: 0.0010379973100498319
step: 450, loss: 0.0018686945550143719
step: 460, loss: 0.019250454381108284
step: 470, loss: 0.0022918167524039745
step: 480, loss: 0.0001944709656527266
step: 490, loss: 0.00010341098095523193
step: 500, loss: 0.000850627722684294
step: 510, loss: 4.927669942844659e-05
step: 520, loss: 0.0002610509400255978
step: 530, loss: 0.05955510586500168
/usr/local/lib/python3.7/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=0.9466292134831461, f1=0.9392523364485983,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008812992833554745
step: 10, loss: 0.0006308113224804401
step: 20, loss: 0.0035843069199472666
step: 30, loss: 5.853292896063067e-05
step: 40, loss: 0.00011043123959098011
step: 50, loss: 2.2429287128034048e-05
step: 60, loss: 2.299577317899093e-05
step: 70, loss: 0.0005116612883284688
step: 80, loss: 6.057493374100886e-05
step: 90, loss: 0.0011630698572844267
step: 100, loss: 0.0038907055277377367
step: 110, loss: 3.985870353062637e-05
step: 120, loss: 3.067931538680568e-05
step: 130, loss: 3.877494600601494e-05
step: 140, loss: 3.929268132196739e-05
step: 150, loss: 2.5363700842717662e-05
step: 160, loss: 6.730639870511368e-05
step: 170, loss: 2.4306173145305365e-05

```

```

step: 180, loss: 1.9751090803765692e-05
step: 190, loss: 0.0005125731695443392
step: 200, loss: 1.5921534213703126e-05
step: 210, loss: 0.0005739190382882953
step: 220, loss: 0.0005363432574085891
step: 230, loss: 1.861865894170478e-05
step: 240, loss: 6.764557474525645e-05
step: 250, loss: 1.545227860333398e-05
step: 260, loss: 1.8223514416604303e-05
step: 270, loss: 0.004238460678607225
step: 280, loss: 2.8721293347189203e-05
step: 290, loss: 0.000376155658159405
step: 300, loss: 0.0012169636320322752
step: 310, loss: 1.8108163203578442e-05
step: 320, loss: 0.0005206165369600058
step: 330, loss: 0.0014170859940350056
step: 340, loss: 8.304492803290486e-05
step: 350, loss: 1.7631304217502475e-05
step: 360, loss: 0.0007572183385491371
step: 370, loss: 0.00037608976708725095
step: 380, loss: 0.019075307995080948
step: 390, loss: 0.0020156302489340305
step: 400, loss: 0.0032607438042759895
step: 410, loss: 0.00017646861670073122
step: 420, loss: 2.9948070732643828e-05
step: 430, loss: 0.00021012680372223258
step: 440, loss: 0.00032252378878183663
step: 450, loss: 0.001024215598590672
step: 460, loss: 0.00018814006762113422
step: 470, loss: 0.003924350254237652
step: 480, loss: 0.0005458163213916123
step: 490, loss: 6.0163332818774506e-05
step: 500, loss: 0.00015391860506497324
step: 510, loss: 0.0014517720555886626
step: 520, loss: 0.0014538965187966824
step: 530, loss: 4.7028577682795e-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=0.9503219871205152, f1=0.9441391941391942,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7499458155944012e-05
step: 10, loss: 0.00017219525761902332
step: 20, loss: 4.1071809391723946e-05
step: 30, loss: 2.31001504289452e-05

```

step: 40, loss: 0.00022112017904873937
step: 50, loss: 0.003380298614501953
step: 60, loss: 3.179003397235647e-05
step: 70, loss: 0.00011800362699432299
step: 80, loss: 6.754929199814796e-05
step: 90, loss: 3.296356226201169e-05
step: 100, loss: 3.2161318813450634e-05
step: 110, loss: 3.2020929211284965e-05
step: 120, loss: 3.1019619200378656e-05
step: 130, loss: 2.0998955733375624e-05
step: 140, loss: 8.791017899056897e-05
step: 150, loss: 0.003807740518823266
step: 160, loss: 6.164106889627874e-05
step: 170, loss: 0.0027608382515609264
step: 180, loss: 3.573531284928322e-05
step: 190, loss: 0.0002241883339593187
step: 200, loss: 3.3222695492440835e-05
step: 210, loss: 2.6098556190845557e-05
step: 220, loss: 2.9509539672289975e-05
step: 230, loss: 0.0013955894391983747
step: 240, loss: 0.0027726064436137676
step: 250, loss: 5.6116732594091445e-05
step: 260, loss: 2.4172784833353944e-05
step: 270, loss: 0.0003150624979753047
step: 280, loss: 3.0851260817144066e-05
step: 290, loss: 1.2133162272220943e-05
step: 300, loss: 1.9628201698651537e-05
step: 310, loss: 3.01513027807232e-05
step: 320, loss: 2.0078914531040937e-05
step: 330, loss: 3.042281423404347e-05
step: 340, loss: 0.0012546611251309514
step: 350, loss: 1.942330527526792e-05
step: 360, loss: 0.0770670473575592
step: 370, loss: 7.491639553336427e-05
step: 380, loss: 0.00430266885086894
step: 390, loss: 2.7003206923836842e-05
step: 400, loss: 2.3763044737279415e-05
step: 410, loss: 2.7256697649136186e-05
step: 420, loss: 1.4971733435231727e-05
step: 430, loss: 0.0001123886468121782
step: 440, loss: 1.6696287275408395e-05
step: 450, loss: 2.6214251192868687e-05
step: 460, loss: 0.08442994952201843
step: 470, loss: 0.00239218445494771
step: 480, loss: 1.2524316844064742e-05
step: 490, loss: 6.015750841470435e-05
step: 500, loss: 0.001820391626097262
step: 510, loss: 5.524485095520504e-05

```

step: 520, loss: 1.3720072274736594e-05
step: 530, loss: 9.538013546261936e-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=0.9477119924457035, f1=0.9450757575757576,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037154420278966427
step: 10, loss: 0.0008719497127458453
step: 20, loss: 1.585835707373917e-05
step: 30, loss: 0.00014942199049983174
step: 40, loss: 0.0001473964366596192
step: 50, loss: 0.0015951166860759258
step: 60, loss: 3.562959318514913e-05
step: 70, loss: 0.00024274663883261383
step: 80, loss: 0.003663405077531934
step: 90, loss: 1.959090877790004e-05
step: 100, loss: 9.678499191068113e-05
step: 110, loss: 9.918327123159543e-05
step: 120, loss: 4.1952469473471865e-05
step: 130, loss: 2.027597656706348e-05
step: 140, loss: 0.00013527962437365204
step: 150, loss: 0.00032609381014481187
step: 160, loss: 0.000692447938490659
step: 170, loss: 0.0010887796524912119
step: 180, loss: 2.832934660546016e-05
step: 190, loss: 0.0009356682421639562
step: 200, loss: 2.4928624043241143e-05
step: 210, loss: 0.00021719779761042446
step: 220, loss: 4.214257933199406e-05
step: 230, loss: 3.9947462937561795e-05
step: 240, loss: 0.0012061366578564048
step: 250, loss: 0.0021969410590827465
step: 260, loss: 8.351437281817198e-05
step: 270, loss: 0.008815444074571133
step: 280, loss: 7.018545875325799e-05
step: 290, loss: 0.0002889694878831506
step: 300, loss: 2.3170710846898146e-05
step: 310, loss: 0.0008389412541873753
step: 320, loss: 2.5604431357351132e-05
step: 330, loss: 2.6195135433226824e-05
step: 340, loss: 4.87511460960377e-05
step: 350, loss: 1.6227100786636584e-05
step: 360, loss: 0.0022247692104429007
step: 370, loss: 0.014658615924417973

```

```

step: 380, loss: 0.0002517457469366491
step: 390, loss: 0.00012442884326446801
step: 400, loss: 0.0017855444457381964
step: 410, loss: 1.6509999113623053e-05
step: 420, loss: 9.111255349125713e-05
step: 430, loss: 0.0010732341324910522
step: 440, loss: 0.0001860000629676506
step: 450, loss: 3.193964221281931e-05
step: 460, loss: 0.03624311834573746
step: 470, loss: 1.293408695346443e-05
step: 480, loss: 2.148683779523708e-05
step: 490, loss: 6.667490379186347e-05
step: 500, loss: 0.0018363037379458547
step: 510, loss: 0.009300176054239273
step: 520, loss: 1.8405884475214407e-05
step: 530, loss: 0.0008268359815701842
/usr/local/lib/python3.7/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=0.9488503050211168, f1=0.9487058823529413,
best_f1=0.9477154969382948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.4401758562598843e-05
step: 10, loss: 0.0018461951985955238
step: 20, loss: 0.0015286379493772984
step: 30, loss: 0.015216663479804993
step: 40, loss: 4.6451070375042036e-05
step: 50, loss: 0.0009819279657676816
step: 60, loss: 9.32088223635219e-05
step: 70, loss: 5.7386434491490945e-05
step: 80, loss: 2.6711819373304024e-05
step: 90, loss: 1.8789460227708332e-05
step: 100, loss: 1.2945248272444587e-05
step: 110, loss: 0.016038939356803894
step: 120, loss: 2.1318568542483263e-05
step: 130, loss: 1.3861646039003972e-05
step: 140, loss: 0.0003929108497686684
step: 150, loss: 1.2058648280799389e-05
step: 160, loss: 2.7672474971041083e-05
step: 170, loss: 8.356612670468166e-05
step: 180, loss: 2.052950185316149e-05
step: 190, loss: 0.0010083713568747044
step: 200, loss: 0.0007061536889523268
step: 210, loss: 6.961877079447731e-05
step: 220, loss: 1.5582601918140426e-05
step: 230, loss: 0.00024820491671562195

```

```

step: 240, loss: 1.2475841685954947e-05
step: 250, loss: 1.7694463167572394e-05
step: 260, loss: 0.00015099119627848268
step: 270, loss: 1.1958052709815092e-05
step: 280, loss: 1.0050760465674102e-05
step: 290, loss: 8.792382868705317e-05
step: 300, loss: 1.9236744265072048e-05
step: 310, loss: 0.03371824324131012
step: 320, loss: 2.024649984377902e-05
step: 330, loss: 1.4956879567762371e-05
step: 340, loss: 8.009768498595804e-05
step: 350, loss: 0.00038502892130054533
step: 360, loss: 2.7527672500582412e-05
step: 370, loss: 4.009467374999076e-05
step: 380, loss: 1.5407562386826612e-05
step: 390, loss: 4.3644013203447685e-05
step: 400, loss: 0.002609915565699339
step: 410, loss: 0.0007372673135250807
step: 420, loss: 3.6043689760845155e-05
step: 430, loss: 1.0281720278726425e-05
step: 440, loss: 1.909530874399934e-05
step: 450, loss: 0.0028191180899739265
step: 460, loss: 0.00036892638308927417
step: 470, loss: 1.4375424143509008e-05
step: 480, loss: 0.003328696358948946
step: 490, loss: 6.681108789052814e-05
step: 500, loss: 0.0011639485601335764
step: 510, loss: 1.5891635484877042e-05
step: 520, loss: 4.1191797208739445e-05
step: 530, loss: 1.059833994077053e-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=0.9493908153701968, f1=0.949343339587242,
best_f1=0.9477154969382948
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 [00:37, 152.32it/s]

load_f1 = 0.9534776600644865

real_f1 = 0.9512308406874129

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

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

4398it [00:32, 133.55it/s]

3.1.4 Amazon-Google - Running the matcher

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

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5041724443435669
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5760746002197266
step: 20, loss: 0.46412545442581177
step: 30, loss: 0.3226455748081207
step: 40, loss: 0.3146035969257355
step: 50, loss: 0.4929715096950531
step: 60, loss: 0.4681403934955597
step: 70, loss: 0.35566043853759766
step: 80, loss: 0.4526368975639343
step: 90, loss: 0.2777218222618103
step: 100, loss: 0.2480766773223877
step: 110, loss: 0.24330869317054749
step: 120, loss: 0.41404128074645996
step: 130, loss: 0.2567008137702942
step: 140, loss: 0.4382891058921814
step: 150, loss: 0.3898755609989166
step: 160, loss: 0.4944058060646057
step: 170, loss: 0.23716744780540466
step: 180, loss: 0.3988436460494995
step: 190, loss: 0.6158481240272522
step: 200, loss: 0.409505158662796
step: 210, loss: 0.511177659034729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3803889751434326
step: 10, loss: 0.2133331596851349
step: 20, loss: 0.48640596866607666
step: 30, loss: 0.5276857614517212
step: 40, loss: 0.4443977475166321
step: 50, loss: 0.21547985076904297
step: 60, loss: 0.3103888928890228
step: 70, loss: 0.430831640958786
step: 80, loss: 0.31116199493408203
step: 90, loss: 0.38683223724365234
step: 100, loss: 0.47620055079460144

```

```

step: 110, loss: 0.3898858428001404
step: 120, loss: 0.23715466260910034
step: 130, loss: 0.1761653870344162
step: 140, loss: 0.2526911497116089
step: 150, loss: 0.4479692578315735
step: 160, loss: 0.16844049096107483
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.6233566999435425
step: 180, loss: 0.3290555477142334
step: 190, loss: 0.3259986639022827
step: 200, loss: 0.15076567232608795
step: 210, loss: 0.30630677938461304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2611106336116791
step: 10, loss: 0.2384774535894394
step: 20, loss: 0.48235952854156494
step: 30, loss: 0.25398707389831543
step: 40, loss: 0.43871983885765076
step: 50, loss: 0.47228479385375977
step: 60, loss: 0.5022034645080566
step: 70, loss: 0.20385396480560303
step: 80, loss: 0.45383796095848083
step: 90, loss: 0.24406981468200684
step: 100, loss: 0.3856942653656006
step: 110, loss: 0.23358485102653503
step: 120, loss: 0.23806773126125336
step: 130, loss: 0.16249431669712067
step: 140, loss: 0.37773597240448
step: 150, loss: 0.3169133961200714
step: 160, loss: 0.21113988757133484
step: 170, loss: 0.38343143463134766
step: 180, loss: 0.2540951669216156
step: 190, loss: 0.16343431174755096
step: 200, loss: 0.2340717315673828
step: 210, loss: 0.2419489026069641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3118242919445038
step: 10, loss: 0.30330562591552734
step: 20, loss: 0.30985352396965027
step: 30, loss: 0.237383633852005
step: 40, loss: 0.23581571877002716
step: 50, loss: 0.23299133777618408
step: 60, loss: 0.4967680275440216
step: 70, loss: 0.25881603360176086
step: 80, loss: 0.23811142146587372
step: 90, loss: 0.4474259912967682
step: 100, loss: 0.3627628982067108
step: 110, loss: 0.5860417485237122
step: 120, loss: 0.3857390284538269
step: 130, loss: 0.6821143627166748
step: 140, loss: 0.49662959575653076
step: 150, loss: 0.3863489329814911
step: 160, loss: 0.3254215121269226
step: 170, loss: 0.18035130202770233
step: 180, loss: 0.08816344290971756
step: 190, loss: 0.16446711122989655
step: 200, loss: 0.317981094121933
step: 210, loss: 0.3802613914012909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38856327533721924
step: 10, loss: 0.32288745045661926
step: 20, loss: 0.32495057582855225
step: 30, loss: 0.23966842889785767
step: 40, loss: 0.4384290277957916
step: 50, loss: 0.38378390669822693
step: 60, loss: 0.38105857372283936
step: 70, loss: 0.24496370553970337
step: 80, loss: 0.45227083563804626
step: 90, loss: 0.4331541359424591
step: 100, loss: 0.2434973120689392
step: 110, loss: 0.15915241837501526
step: 120, loss: 0.22988271713256836
step: 130, loss: 0.3169121742248535
step: 140, loss: 0.5308547616004944
step: 150, loss: 0.31218358874320984

```

```

step: 160, loss: 0.24974587559700012
step: 170, loss: 0.3843604028224945
step: 180, loss: 0.24003081023693085
step: 190, loss: 0.5310019254684448
step: 200, loss: 0.38769400119781494
step: 210, loss: 0.2459791749715805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16266445815563202
step: 10, loss: 0.38178667426109314
step: 20, loss: 0.30915138125419617
step: 30, loss: 0.2449016124010086
step: 40, loss: 0.31889083981513977
step: 50, loss: 0.6135328412055969
step: 60, loss: 0.3136651813983917
step: 70, loss: 0.4336855113506317
step: 80, loss: 0.311686247587204
step: 90, loss: 0.31193703413009644
step: 100, loss: 0.3157855272293091
step: 110, loss: 0.3762960135936737
step: 120, loss: 0.24778659641742706
step: 130, loss: 0.2360650897026062
step: 140, loss: 0.37968653440475464
step: 150, loss: 0.18871928751468658
step: 160, loss: 0.16579945385456085
step: 170, loss: 0.4601098895072937
step: 180, loss: 0.3770511746406555
step: 190, loss: 0.45443665981292725
step: 200, loss: 0.30551791191101074
step: 210, loss: 0.3796556890010834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4500158131122589
step: 10, loss: 0.31283923983573914
step: 20, loss: 0.4407801330089569
step: 30, loss: 0.24646779894828796

```

```

step: 40, loss: 0.24704232811927795
step: 50, loss: 0.37914079427719116
step: 60, loss: 0.3170207440853119
step: 70, loss: 0.2524263262748718
step: 80, loss: 0.5110667943954468
step: 90, loss: 0.37351059913635254
step: 100, loss: 0.4387146532535553
step: 110, loss: 0.30531415343284607
step: 120, loss: 0.31798070669174194
step: 130, loss: 0.47109851241111755
step: 140, loss: 0.3137548863887787
step: 150, loss: 0.3130759596824646
step: 160, loss: 0.5521305203437805
step: 170, loss: 0.618061363697052
step: 180, loss: 0.23477229475975037
step: 190, loss: 0.23835596442222595
step: 200, loss: 0.3125579357147217
step: 210, loss: 0.31427887082099915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6080752611160278
step: 10, loss: 0.31781715154647827
step: 20, loss: 0.24929334223270416
step: 30, loss: 0.24800190329551697
step: 40, loss: 0.2410147488117218
step: 50, loss: 0.16584770381450653
step: 60, loss: 0.16983480751514435
step: 70, loss: 0.45943328738212585
step: 80, loss: 0.30769720673561096
step: 90, loss: 0.3851092457771301
step: 100, loss: 0.6102311611175537
step: 110, loss: 0.37461456656455994
step: 120, loss: 0.30152660608291626
step: 130, loss: 0.1708136647939682
step: 140, loss: 0.3168551027774811
step: 150, loss: 0.4711753726005554
step: 160, loss: 0.44375962018966675
step: 170, loss: 0.5304307341575623
step: 180, loss: 0.3129090666770935
step: 190, loss: 0.2296237349510193
step: 200, loss: 0.45537886023521423
step: 210, loss: 0.38395410776138306

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5961818099021912
step: 10, loss: 0.28420498967170715
step: 20, loss: 0.38480904698371887
step: 30, loss: 0.17202965915203094
step: 40, loss: 0.16779232025146484
step: 50, loss: 0.45096951723098755
step: 60, loss: 0.17190766334533691
step: 70, loss: 0.38450485467910767
step: 80, loss: 0.30622780323028564
step: 90, loss: 0.32051917910575867
step: 100, loss: 0.46216824650764465
step: 110, loss: 0.5290734767913818
step: 120, loss: 0.38639986515045166
step: 130, loss: 0.24259072542190552
step: 140, loss: 0.5970783233642578
step: 150, loss: 0.11306048929691315
step: 160, loss: 0.3782874047756195
step: 170, loss: 0.3169347941875458
step: 180, loss: 0.4573005139827728
step: 190, loss: 0.31084802746772766
step: 200, loss: 0.31131434440612793
step: 210, loss: 0.30659669637680054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.306844562292099
step: 10, loss: 0.3826238214969635
step: 20, loss: 0.25994354486465454
step: 30, loss: 0.17355534434318542
step: 40, loss: 0.3121204972267151
step: 50, loss: 0.5112889409065247
step: 60, loss: 0.24636225402355194
step: 70, loss: 0.3845391273498535
step: 80, loss: 0.1636137068271637
step: 90, loss: 0.38423317670822144

```



```

step: 100, loss: 0.46097037196159363
step: 110, loss: 0.17404741048812866
step: 120, loss: 0.4508204162120819
step: 130, loss: 0.24040956795215607
step: 140, loss: 0.3726884722709656
step: 150, loss: 0.24766044318675995
step: 160, loss: 0.3831023573875427
step: 170, loss: 0.24886077642440796
step: 180, loss: 0.3160536289215088
step: 190, loss: 0.24149826169013977
step: 200, loss: 0.6418158411979675
step: 210, loss: 0.2511022984981537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3166930079460144
step: 10, loss: 0.31459182500839233
step: 20, loss: 0.3901274800300598
step: 30, loss: 0.16756044328212738
step: 40, loss: 0.4468727707862854
step: 50, loss: 0.4329097270965576
step: 60, loss: 0.44220489263534546
step: 70, loss: 0.17682139575481415
step: 80, loss: 0.5222845673561096
step: 90, loss: 0.3908003568649292
step: 100, loss: 0.5137608647346497
step: 110, loss: 0.44887274503707886
step: 120, loss: 0.44418230652809143
step: 130, loss: 0.24418407678604126
step: 140, loss: 0.31310388445854187
step: 150, loss: 0.31894001364707947
step: 160, loss: 0.1715497076511383
step: 170, loss: 0.3132527768611908
step: 180, loss: 0.23998220264911652
step: 190, loss: 0.4468246400356293
step: 200, loss: 0.1798144280910492
step: 210, loss: 0.3758852779865265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24808506667613983
step: 10, loss: 0.30906951427459717
step: 20, loss: 0.46283408999443054
step: 30, loss: 0.3114384710788727
step: 40, loss: 0.17242968082427979
step: 50, loss: 0.45090898871421814
step: 60, loss: 0.11018909513950348
step: 70, loss: 0.44754862785339355
step: 80, loss: 0.31638020277023315
step: 90, loss: 0.390511155128479
step: 100, loss: 0.12397093325853348
step: 110, loss: 0.2420874536037445
step: 120, loss: 0.2459345906972885
step: 130, loss: 0.4610857665538788
step: 140, loss: 0.3796297609806061
step: 150, loss: 0.0994202271103859
step: 160, loss: 0.24181769788265228
step: 170, loss: 0.2402547299861908
step: 180, loss: 0.3151618242263794
step: 190, loss: 0.3818625509738922
step: 200, loss: 0.17403188347816467
step: 210, loss: 0.3139626979827881
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3114413321018219
step: 10, loss: 0.2374451905488968
step: 20, loss: 0.31465575098991394
step: 30, loss: 0.24646402895450592
step: 40, loss: 0.25011345744132996
step: 50, loss: 0.24861694872379303
step: 60, loss: 0.5764137506484985
step: 70, loss: 0.4433172643184662
step: 80, loss: 0.38308677077293396
step: 90, loss: 0.24670444428920746
step: 100, loss: 0.3796460032463074
step: 110, loss: 0.24519012868404388
step: 120, loss: 0.30535081028938293
step: 130, loss: 0.31033751368522644
step: 140, loss: 0.16543924808502197
step: 150, loss: 0.31239476799964905

```

```

step: 160, loss: 0.5401768684387207
step: 170, loss: 0.2417943775653839
step: 180, loss: 0.2387385219335556
step: 190, loss: 0.24519082903862
step: 200, loss: 0.38060498237609863
step: 210, loss: 0.37421509623527527
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24273380637168884
step: 10, loss: 0.3824155330657959
step: 20, loss: 0.5109807252883911
step: 30, loss: 0.24638235569000244
step: 40, loss: 0.32907241582870483
step: 50, loss: 0.25099894404411316
step: 60, loss: 0.379548043012619
step: 70, loss: 0.24632269144058228
step: 80, loss: 0.3163859248161316
step: 90, loss: 0.17871858179569244
step: 100, loss: 0.2423788458108902
step: 110, loss: 0.44332605600357056
step: 120, loss: 0.24852672219276428
step: 130, loss: 0.3109099864959717
step: 140, loss: 0.31016817688941956
step: 150, loss: 0.31566542387008667
step: 160, loss: 0.1778554618358612
step: 170, loss: 0.45246753096580505
step: 180, loss: 0.2411431223154068
step: 190, loss: 0.3858364224433899
step: 200, loss: 0.17004847526550293
step: 210, loss: 0.4496334493160248
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2425592690706253
step: 10, loss: 0.2450638860464096
step: 20, loss: 0.5192540884017944
step: 30, loss: 0.24554231762886047

```

```

step: 40, loss: 0.38560301065444946
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.1735723465681076
step: 60, loss: 0.30145877599716187
step: 70, loss: 0.45520392060279846
step: 80, loss: 0.2399044781923294
step: 90, loss: 0.452409565448761
step: 100, loss: 0.3116222321987152
step: 110, loss: 0.3115505576133728
step: 120, loss: 0.45081785321235657
step: 130, loss: 0.17127947509288788
step: 140, loss: 0.3101363778114319
step: 150, loss: 0.5316551327705383
step: 160, loss: 0.37952762842178345
step: 170, loss: 0.3787423372268677
step: 180, loss: 0.308510959148407
step: 190, loss: 0.31464287638664246
step: 200, loss: 0.31146353483200073
step: 210, loss: 0.4577195644378662
/usr/local/lib/python3.7/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=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
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 [00:09, 230.31it/s]
load_f1 = 0.18519984170953702
real_f1 = 0.18519984170953702
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 [00:34, 129.06it/s]

```

3.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

```
--checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

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

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

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

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

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

step: 0, loss: 0.46847236156463623

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

step: 10, loss: 0.4115792512893677

step: 20, loss: 0.25278806686401367

step: 30, loss: 0.3987118601799011

step: 40, loss: 0.24234120547771454

step: 50, loss: 0.3330713212490082

step: 60, loss: 0.49866411089897156

step: 70, loss: 0.4431764483451843

```

step: 80, loss: 0.15188930928707123
step: 90, loss: 0.31428030133247375
step: 100, loss: 0.4199192225933075
step: 110, loss: 0.25078678131103516
step: 120, loss: 0.31907281279563904
step: 130, loss: 0.3459508717060089
step: 140, loss: 0.17237292230129242
step: 150, loss: 0.32132771611213684
step: 160, loss: 0.22802986204624176
step: 170, loss: 0.365928590297699
step: 180, loss: 0.17943677306175232
step: 190, loss: 0.15548226237297058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39442339539527893
step: 10, loss: 0.34057876467704773
step: 20, loss: 0.6717296242713928
step: 30, loss: 0.25412312150001526
step: 40, loss: 0.5475385785102844
step: 50, loss: 0.33722439408302307
step: 60, loss: 0.4413849413394928
step: 70, loss: 0.3136940002441406
step: 80, loss: 0.1808326542377472
step: 90, loss: 0.2864428758621216
step: 100, loss: 0.15660053491592407
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.20496827363967896
step: 120, loss: 0.17856065928936005
step: 130, loss: 0.1341363489627838
step: 140, loss: 0.19769425690174103
step: 150, loss: 0.29242628812789917
step: 160, loss: 0.2163134664297104
step: 170, loss: 0.04637683928012848
step: 180, loss: 0.036936867982149124
step: 190, loss: 0.09708469361066818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6975609756097562, f1=0.7004830917874396,
best_f1=0.7004830917874396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18086282908916473
step: 10, loss: 0.18451061844825745
step: 20, loss: 0.20310895144939423
step: 30, loss: 0.13614994287490845
step: 40, loss: 0.1388663947582245
step: 50, loss: 0.07965706288814545
step: 60, loss: 0.12410826981067657
step: 70, loss: 0.1578894853591919
step: 80, loss: 0.05592035502195358
step: 90, loss: 0.1582416594028473
step: 100, loss: 0.16150324046611786
step: 110, loss: 0.30355599522590637
step: 120, loss: 0.20854546129703522
step: 130, loss: 0.03865088149905205
step: 140, loss: 0.09753090143203735
step: 150, loss: 0.1500217616558075
step: 160, loss: 0.11065662652254105
step: 170, loss: 0.3195893168449402
step: 180, loss: 0.13831143081188202
step: 190, loss: 0.09389946609735489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8241758241758241, f1=0.8176795580110496,
best_f1=0.8176795580110496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01914484240114689
step: 10, loss: 0.11206596344709396
step: 20, loss: 0.03542495518922806
step: 30, loss: 0.013778485357761383
step: 40, loss: 0.040788255631923676
step: 50, loss: 0.03343485668301582
step: 60, loss: 0.151172935962677

```



```

step: 70, loss: 0.11115764081478119
step: 80, loss: 0.023004934191703796
step: 90, loss: 0.008660678751766682
step: 100, loss: 0.07500115782022476
step: 110, loss: 0.09339321404695511
step: 120, loss: 0.029536087065935135
step: 130, loss: 0.03077978640794754
step: 140, loss: 0.22203421592712402
step: 150, loss: 0.013278120197355747
step: 160, loss: 0.021513523533940315
step: 170, loss: 0.037084586918354034
step: 180, loss: 0.054670657962560654
step: 190, loss: 0.015557521022856236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8262108262108262, f1=0.8158640226628895,
best_f1=0.8158640226628895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2945418357849121
step: 10, loss: 0.1293381005525589
step: 20, loss: 0.03902410343289375
step: 30, loss: 0.010798677802085876
step: 40, loss: 0.03664006292819977
step: 50, loss: 0.00929232593625784
step: 60, loss: 0.0028067012317478657
step: 70, loss: 0.017530910670757294
step: 80, loss: 0.017480066046118736
step: 90, loss: 0.20033012330532074
step: 100, loss: 0.12418468296527863
step: 110, loss: 0.30116522312164307
step: 120, loss: 0.0363299697637558
step: 130, loss: 0.32366645336151123
step: 140, loss: 0.03341720253229141
step: 150, loss: 0.06713426113128662
step: 160, loss: 0.10145779699087143
step: 170, loss: 0.01770769990980625
step: 180, loss: 0.01584768109023571
step: 190, loss: 0.1548897922039032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8410256410256409, f1=0.7958656330749353,
best_f1=0.7958656330749353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026647381484508514
step: 10, loss: 0.02006879262626171
step: 20, loss: 0.023508938029408455
step: 30, loss: 0.03958183154463768
step: 40, loss: 0.011271040886640549
step: 50, loss: 0.009938502684235573
step: 60, loss: 0.011587628163397312
step: 70, loss: 0.007776594255119562
step: 80, loss: 0.006697839591652155
step: 90, loss: 0.006057559046894312
step: 100, loss: 0.030241243541240692
step: 110, loss: 0.010850438848137856
step: 120, loss: 0.018137061968445778
step: 130, loss: 0.15069922804832458
step: 140, loss: 0.03271099925041199
step: 150, loss: 0.10286326706409454
step: 160, loss: 0.15831461548805237
step: 170, loss: 0.30506443977355957
step: 180, loss: 0.04873563349246979
step: 190, loss: 0.011439301073551178
/usr/local/lib/python3.7/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=0.8365650969529086, f1=0.8328767123287673,
best_f1=0.7958656330749353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014616257511079311
step: 10, loss: 0.002826160052791238
step: 20, loss: 0.0013083535013720393
step: 30, loss: 0.07986722886562347
step: 40, loss: 0.007554213982075453
step: 50, loss: 0.0028622036334127188
step: 60, loss: 0.05793379992246628
step: 70, loss: 0.0034682455006986856
step: 80, loss: 0.007118529174476862
step: 90, loss: 0.004257110878825188
step: 100, loss: 0.2275584638118744

```

```

step: 110, loss: 0.08180779963731766
step: 120, loss: 0.03298849239945412
step: 130, loss: 0.017613181844353676
step: 140, loss: 0.03495578095316887
step: 150, loss: 0.017190175130963326
step: 160, loss: 0.2438485026359558
step: 170, loss: 0.0700046494603157
step: 180, loss: 0.011344973929226398
step: 190, loss: 0.05285221338272095
/usr/local/lib/python3.7/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=0.8191489361702128, f1=0.8333333333333334,
best_f1=0.7958656330749353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012341393157839775
step: 10, loss: 0.0025229353923350573
step: 20, loss: 0.0011171952355653048
step: 30, loss: 0.002647532382979989
step: 40, loss: 0.07916822284460068
step: 50, loss: 0.03030943125486374
step: 60, loss: 0.009069283492863178
step: 70, loss: 0.001902374206110835
step: 80, loss: 0.01921258494257927
step: 90, loss: 0.002786843106150627
step: 100, loss: 0.0009462088928557932
step: 110, loss: 0.05603628233075142
step: 120, loss: 0.0017545379232615232
step: 130, loss: 0.0025458973832428455
step: 140, loss: 0.0016760111320763826
step: 150, loss: 0.00809592753648758
step: 160, loss: 0.00162902707234025
step: 170, loss: 0.0010482539655640721
step: 180, loss: 0.18499746918678284
step: 190, loss: 0.00129341846331954
/usr/local/lib/python3.7/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=0.8168316831683167, f1=0.8418367346938775,
best_f1=0.7958656330749353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023974638897925615
step: 10, loss: 0.0006520007736980915
step: 20, loss: 0.0033706151880323887

```

```

step: 30, loss: 0.006080552004277706
step: 40, loss: 0.020642917603254318
step: 50, loss: 0.02332977205514908
step: 60, loss: 0.0021090267691761255
step: 70, loss: 0.0008931256597861648
step: 80, loss: 0.022474534809589386
step: 90, loss: 0.005330640822649002
step: 100, loss: 0.2595476508140564
step: 110, loss: 0.0019356028642505407
step: 120, loss: 0.039438147097826004
step: 130, loss: 0.0027646494563668966
step: 140, loss: 0.010507187806069851
step: 150, loss: 0.003366640768945217
step: 160, loss: 0.0012045946205034852
step: 170, loss: 0.0011147403856739402
step: 180, loss: 0.0038143189158290625
step: 190, loss: 0.0006192188593558967
/usr/local/lib/python3.7/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=0.8232323232323232, f1=0.835509138381201,
best_f1=0.7958656330749353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010528888087719679
step: 10, loss: 0.00039832296897657216
step: 20, loss: 0.14058014750480652
step: 30, loss: 0.11279873549938202
step: 40, loss: 0.005860275123268366
step: 50, loss: 0.0009736421634443104
step: 60, loss: 0.0007933145971037447
step: 70, loss: 0.001491026021540165
step: 80, loss: 0.001766063505783677
step: 90, loss: 0.001974229235202074
step: 100, loss: 0.0016441134503111243
step: 110, loss: 0.00393324252218008
step: 120, loss: 0.0013563059037551284
step: 130, loss: 0.0007615129579789937
step: 140, loss: 0.001618041074834764
step: 150, loss: 0.007270730100572109
step: 160, loss: 0.001909139333292842
step: 170, loss: 0.0007197754457592964
step: 180, loss: 0.002033962169662118
step: 190, loss: 0.01179587934166193
/usr/local/lib/python3.7/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=0.8244680851063829, f1=0.8364611260053619,
best_f1=0.7958656330749353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005586084444075823
step: 10, loss: 0.0004581343091558665
step: 20, loss: 0.0024623223580420017
step: 30, loss: 0.0019119747448712587
step: 40, loss: 0.0012783717829734087
step: 50, loss: 0.000694502901751548
step: 60, loss: 0.0018231788417324424
step: 70, loss: 0.0007630238542333245
step: 80, loss: 0.053856294602155685
step: 90, loss: 0.00028822096646763384
step: 100, loss: 0.0008112555951811373
step: 110, loss: 0.0900740772485733
step: 120, loss: 0.000719624396879226
step: 130, loss: 0.0015355248469859362
step: 140, loss: 0.012341001071035862
step: 150, loss: 0.0032894352916628122
step: 160, loss: 0.003829773049801588
step: 170, loss: 0.000639323377981782
step: 180, loss: 0.002005456481128931
step: 190, loss: 0.0003895942645613104
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8431876606683805, f1=0.84375, best_f1=0.84375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003988250973634422
step: 10, loss: 0.0033305136021226645
step: 20, loss: 0.0009184475056827068
step: 30, loss: 0.001074274186976254
step: 40, loss: 0.0030125800985842943
step: 50, loss: 0.0014122631400823593
step: 60, loss: 0.11075753718614578
step: 70, loss: 0.00660345796495676
step: 80, loss: 0.0013840642059221864
step: 90, loss: 0.0008655470446683466
step: 100, loss: 0.0009214226738549769
step: 110, loss: 0.006240652408450842

```

```

step: 120, loss: 0.0004540653608273715
step: 130, loss: 0.003048297483474016
step: 140, loss: 0.00026682051247917116
step: 150, loss: 0.0013332534581422806
step: 160, loss: 0.0007127244607545435
step: 170, loss: 0.00048141900333575904
step: 180, loss: 0.00023646761837881058
step: 190, loss: 0.00038997404044494033
/usr/local/lib/python3.7/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=0.8364611260053619, f1=0.8406593406593407, best_f1=0.84375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006560764741152525
step: 10, loss: 0.0017428739229217172
step: 20, loss: 0.00019283870642539114
step: 30, loss: 0.0019952713046222925
step: 40, loss: 0.004327396862208843
step: 50, loss: 0.001434607314877212
step: 60, loss: 0.0700850859284401
step: 70, loss: 0.0009389185579493642
step: 80, loss: 0.000752704800106585
step: 90, loss: 0.0002151427761418745
step: 100, loss: 0.0003272911417298019
step: 110, loss: 0.004590901546180248
step: 120, loss: 0.00034889098606072366
step: 130, loss: 0.0004872263816650957
step: 140, loss: 0.009172727353870869
step: 150, loss: 0.00031474410207010806
step: 160, loss: 0.0011071108747273684
step: 170, loss: 0.00022637276560999453
step: 180, loss: 0.0004288263153284788
step: 190, loss: 0.11701346188783646
/usr/local/lib/python3.7/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=0.8409703504043127, f1=0.8324022346368716, best_f1=0.84375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016918780747801065
step: 10, loss: 0.01138338539749384
step: 20, loss: 0.014475353062152863
step: 30, loss: 0.0003854901297017932
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.000356510397978127

```

```

step: 50, loss: 0.0003567979147192091
step: 60, loss: 0.0003618735645432025
step: 70, loss: 0.0003009666397701949
step: 80, loss: 0.0007518581696785986
step: 90, loss: 0.0013826419599354267
step: 100, loss: 0.00011932936467928812
step: 110, loss: 0.00011864330008393154
step: 120, loss: 0.0008615721017122269
step: 130, loss: 0.00019052106654271483
step: 140, loss: 0.0002612588868942112
step: 150, loss: 0.00040054417331703007
step: 160, loss: 0.00016518999473191798
step: 170, loss: 0.000199367044842802
step: 180, loss: 0.00039204375934787095
step: 190, loss: 0.0005193006945773959
/usr/local/lib/python3.7/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=0.8409703504043127, f1=0.8509485094850948, best_f1=0.84375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01730506680905819
step: 10, loss: 0.00013116448826622218
step: 20, loss: 0.003324739634990692
step: 30, loss: 0.00019445974612608552
step: 40, loss: 0.00044488071580417454
step: 50, loss: 0.01978858932852745
step: 60, loss: 0.0010715678799897432
step: 70, loss: 0.0007549096480943263
step: 80, loss: 0.0002680705802049488
step: 90, loss: 0.0038753501139581203
step: 100, loss: 0.008258750662207603
step: 110, loss: 0.0003865629551000893
step: 120, loss: 0.0002965984749607742
step: 130, loss: 0.00028002492035739124
step: 140, loss: 0.00022580068616662174
step: 150, loss: 0.0001387081720167771
step: 160, loss: 0.00014912104234099388
step: 170, loss: 0.00024052814114838839
step: 180, loss: 0.0005970289348624647
step: 190, loss: 0.0005443572299554944
/usr/local/lib/python3.7/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=0.8342245989304813, f1=0.8509485094850948, best_f1=0.84375
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.")
```

2049it [00:14, 139.18it/s]

load_f1 = 0.7916666666666666

real_f1 = 0.7857142857142857

733it [00:00, 2583.07it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

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

4398it [00:33, 131.43it/s]

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

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



```

--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

```

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

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

Step 3: run the matcher

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5249601602554321
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5497928261756897
step: 20, loss: 0.351473331451416
step: 30, loss: 0.3982040584087372
step: 40, loss: 0.6531309485435486
step: 50, loss: 0.3058395981788635
step: 60, loss: 0.5454572439193726
step: 70, loss: 0.32871755957603455
step: 80, loss: 0.23036383092403412
step: 90, loss: 0.2341386079788208
step: 100, loss: 0.151798814535141
step: 110, loss: 0.3993971347808838
step: 120, loss: 0.3290098011493683
step: 130, loss: 0.28678056597709656
step: 140, loss: 0.37784543633461
step: 150, loss: 0.30124741792678833
step: 160, loss: 0.4085273742675781
step: 170, loss: 0.32289448380470276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24733475479744135, f1=0.2648401826484018,
best_f1=0.2648401826484018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30023038387298584
step: 10, loss: 0.4237901568412781

```

```

step: 20, loss: 0.24592894315719604
step: 30, loss: 0.29630425572395325
step: 40, loss: 0.2149616926908493
step: 50, loss: 0.16563168168067932
step: 60, loss: 0.08697736263275146
step: 70, loss: 0.2655879259109497
step: 80, loss: 0.3065975308418274
step: 90, loss: 0.09868501126766205
step: 100, loss: 0.23029345273971558
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.2423168122768402
step: 120, loss: 0.16240696609020233
step: 130, loss: 0.21638941764831543
step: 140, loss: 0.18939340114593506
step: 150, loss: 0.11011567711830139
step: 160, loss: 0.15785230696201324
step: 170, loss: 0.02424352429807186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7927107061503416, f1=0.7775377969762419,
best_f1=0.7775377969762419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24926508963108063
step: 10, loss: 0.0770546942949295
step: 20, loss: 0.1011771485209465
step: 30, loss: 0.11170203238725662
step: 40, loss: 0.13045582175254822
step: 50, loss: 0.20202551782131195
step: 60, loss: 0.042885877192020416
step: 70, loss: 0.14096154272556305
step: 80, loss: 0.026692185550928116
step: 90, loss: 0.43607962131500244
step: 100, loss: 0.01649686135351658
step: 110, loss: 0.02083825320005417
step: 120, loss: 0.06812901049852371
step: 130, loss: 0.2071593850851059
step: 140, loss: 0.10489823669195175
step: 150, loss: 0.012388534843921661
step: 160, loss: 0.05916867405176163
step: 170, loss: 0.03635726496577263
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965517, f1=0.8510638297872338,
best_f1=0.8510638297872338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04138745367527008
step: 10, loss: 0.0050291339866817
step: 20, loss: 0.12619049847126007
step: 30, loss: 0.1932460516691208
step: 40, loss: 0.02667800895869732
step: 50, loss: 0.029154254123568535
step: 60, loss: 0.1279328316450119
step: 70, loss: 0.0022752524819225073
step: 80, loss: 0.124204121530056
step: 90, loss: 0.13153356313705444
step: 100, loss: 0.0752420574426651
step: 110, loss: 0.22493118047714233
step: 120, loss: 0.10672719031572342
step: 130, loss: 0.03223070129752159
step: 140, loss: 0.07547540962696075
step: 150, loss: 0.07491065561771393
step: 160, loss: 0.003172406228259206
step: 170, loss: 0.14800181984901428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8300970873786409, f1=0.859122401847575,
best_f1=0.859122401847575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008608822710812092
step: 10, loss: 0.06692766398191452
step: 20, loss: 0.07176415622234344
step: 30, loss: 0.01643216609954834
step: 40, loss: 0.10107723623514175
step: 50, loss: 0.013041311874985695
step: 60, loss: 0.05108681693673134

```

```

step: 70, loss: 0.007237757556140423
step: 80, loss: 0.001257374999113381
step: 90, loss: 0.021047376096248627
step: 100, loss: 0.029879016801714897
step: 110, loss: 0.04678644239902496
step: 120, loss: 0.017711380496621132
step: 130, loss: 0.004139070399105549
step: 140, loss: 0.01957232505083084
step: 150, loss: 0.02404184639453888
step: 160, loss: 0.0026144927833229303
step: 170, loss: 0.0063479505479335785
/usr/local/lib/python3.7/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=0.8207792207792208, f1=0.8676470588235294,
best_f1=0.859122401847575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13430653512477875
step: 10, loss: 0.027685604989528656
step: 20, loss: 0.0718703493475914
step: 30, loss: 0.004349185153841972
step: 40, loss: 0.0008456888608634472
step: 50, loss: 0.0033081634901463985
step: 60, loss: 0.03153763338923454
step: 70, loss: 0.007740318309515715
step: 80, loss: 0.004386541433632374
step: 90, loss: 0.09875376522541046
step: 100, loss: 0.027656545862555504
step: 110, loss: 0.027118466794490814
step: 120, loss: 0.018277736380696297
step: 130, loss: 0.005712701473385096
step: 140, loss: 0.010577969253063202
step: 150, loss: 0.09469900280237198
step: 160, loss: 0.08332762122154236
step: 170, loss: 0.005218413658440113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8522167487684729, f1=0.8738317757009346,
best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04046444222331047
step: 10, loss: 0.005959186237305403
step: 20, loss: 0.01164967380464077
step: 30, loss: 0.00802281592041254
step: 40, loss: 0.015076511539518833
step: 50, loss: 0.0004776997084263712
step: 60, loss: 0.027882792055606842
step: 70, loss: 0.006263351533561945
step: 80, loss: 0.008603023365139961
step: 90, loss: 0.0727853998541832
step: 100, loss: 0.0007897292380221188
step: 110, loss: 0.013930447399616241
step: 120, loss: 0.005298496223986149
step: 130, loss: 0.0020485869608819485
step: 140, loss: 0.0010326574556529522
step: 150, loss: 0.06305106729269028
step: 160, loss: 0.004041845910251141
step: 170, loss: 0.018704991787672043
/usr/local/lib/python3.7/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=0.8259740259740259, f1=0.8676470588235294,
best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05324633792042732
step: 10, loss: 0.004341420251876116
step: 20, loss: 0.0014787534018978477
step: 30, loss: 0.0003672220336738974
step: 40, loss: 0.0004409459070302546
step: 50, loss: 0.009924323298037052
step: 60, loss: 0.0015466201584786177
step: 70, loss: 0.005059476010501385
step: 80, loss: 0.01227517519146204
step: 90, loss: 0.052096568048000336
step: 100, loss: 0.0002523709845263511
step: 110, loss: 0.24808524549007416
step: 120, loss: 0.03241337090730667
step: 130, loss: 0.0017387925181537867
step: 140, loss: 0.007868831045925617
step: 150, loss: 0.006447015330195427
step: 160, loss: 0.06742282211780548
step: 170, loss: 0.08587995171546936
/usr/local/lib/python3.7/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=0.828125, f1=0.8710462287104622, best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007912680506706238
step: 10, loss: 0.007838549092411995
step: 20, loss: 0.0073853530921041965
step: 30, loss: 0.01232126448303461
step: 40, loss: 0.003198971739038825
step: 50, loss: 0.002621992025524378
step: 60, loss: 0.02086629532277584
step: 70, loss: 0.0022360251750797033
step: 80, loss: 0.0011136432876810431
step: 90, loss: 0.0009522187174297869
step: 100, loss: 0.02453494258224964
step: 110, loss: 0.12995027005672455
step: 120, loss: 0.04685494303703308
step: 130, loss: 0.0011424445547163486
step: 140, loss: 0.0004720531578641385
step: 150, loss: 0.04525764286518097
step: 160, loss: 0.0025086114183068275
step: 170, loss: 0.010014726780354977
/usr/local/lib/python3.7/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=0.8298969072164948, f1=0.8641975308641976,
best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04038283973932266
step: 10, loss: 0.0005129383644089103
step: 20, loss: 0.053680580109357834
step: 30, loss: 0.07247604429721832
step: 40, loss: 0.002271861769258976
step: 50, loss: 0.029350392520427704
step: 60, loss: 0.01180072408169508
step: 70, loss: 0.0011072501074522734
step: 80, loss: 0.04036116227507591
step: 90, loss: 0.0002308545372216031
step: 100, loss: 0.008553266525268555
step: 110, loss: 0.09055469930171967
step: 120, loss: 0.004424441140145063
step: 130, loss: 0.015338076278567314
step: 140, loss: 0.0004361289902590215
step: 150, loss: 0.00014488125452771783
step: 160, loss: 0.0005711041740141809
step: 170, loss: 0.04130905494093895

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8549618320610687, f1=0.8782816229116946,
best_f1=0.8782816229116946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01620177924633026
step: 10, loss: 0.00343704828992486
step: 20, loss: 0.00013123148528393358
step: 30, loss: 0.005841887556016445
step: 40, loss: 0.00013685256999451667
step: 50, loss: 0.0767478495836258
step: 60, loss: 0.0087158577516675
step: 70, loss: 0.0004456837777979672
step: 80, loss: 0.029096094891428947
step: 90, loss: 0.012621023692190647
step: 100, loss: 0.0019184760749340057
step: 110, loss: 0.00019952333241235465
step: 120, loss: 0.000598473590798676
step: 130, loss: 0.015750307589769363
step: 140, loss: 0.0002816115156747401
step: 150, loss: 0.00010800376912811771
step: 160, loss: 0.007136739324778318
step: 170, loss: 0.0058453381061553955
/usr/local/lib/python3.7/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=0.8513853904282115, f1=0.8646080760095013,
best_f1=0.8782816229116946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007168698124587536
step: 10, loss: 0.00011836887279059738
step: 20, loss: 0.00012179333134554327
step: 30, loss: 0.05279269441962242
step: 40, loss: 0.00011945598089369014
step: 50, loss: 0.0001018785042106174
step: 60, loss: 0.0003291392349638045
step: 70, loss: 9.893744572764263e-05
step: 80, loss: 0.00013888526882510632
step: 90, loss: 0.007895654067397118

```



```

step: 100, loss: 0.0008704950450919569
step: 110, loss: 0.028484143316745758
step: 120, loss: 0.04059790074825287
step: 130, loss: 0.0049604992382228374
step: 140, loss: 0.0001251690846402198
step: 150, loss: 0.008855103515088558
step: 160, loss: 0.00540638854727149
step: 170, loss: 0.007926074787974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.858611825192802, f1=0.8695652173913043,
best_f1=0.8695652173913043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007048183120787144
step: 10, loss: 0.0002149652864318341
step: 20, loss: 0.00030577636789530516
step: 30, loss: 0.02018013782799244
step: 40, loss: 0.00016009205137379467
step: 50, loss: 0.005400970112532377
step: 60, loss: 0.00012462426093406975
step: 70, loss: 0.04960281401872635
step: 80, loss: 0.001833549002185464
step: 90, loss: 0.001456701778806746
step: 100, loss: 0.049019020050764084
step: 110, loss: 0.0006198078626766801
step: 120, loss: 0.030600933358073235
step: 130, loss: 7.855131843825802e-05
step: 140, loss: 0.002261360874399543
step: 150, loss: 0.00012189291010145098
step: 160, loss: 0.0002370007277932018
step: 170, loss: 8.613472891738638e-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=0.8324324324324324, f1=0.8706467661691543,
best_f1=0.8695652173913043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011060868389904499
step: 10, loss: 0.00023988615430425853

```

```

step: 20, loss: 0.007028400897979736
step: 30, loss: 0.014924558810889721
step: 40, loss: 0.00012607777898665518
step: 50, loss: 0.00024511016090400517
step: 60, loss: 0.001121705980040133
step: 70, loss: 0.00014673813711851835
step: 80, loss: 0.00022274126240517944
step: 90, loss: 7.640466355951503e-05
step: 100, loss: 0.00015337896184064448
step: 110, loss: 0.04126952961087227
step: 120, loss: 0.0002180957089876756
step: 130, loss: 0.06163746118545532
step: 140, loss: 0.07146011292934418
step: 150, loss: 0.02199455164372921
step: 160, loss: 8.232030813815072e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.015758957713842392
/usr/local/lib/python3.7/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=0.8321167883211679, f1=0.8577981651376146,
best_f1=0.8695652173913043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.385643791873008e-05
step: 10, loss: 0.00016293123189825565
step: 20, loss: 0.00010364706395193934
step: 30, loss: 0.00013828417286276817
step: 40, loss: 6.556610605912283e-05
step: 50, loss: 8.666322537465021e-05
step: 60, loss: 0.003160474356263876
step: 70, loss: 0.04081973060965538
step: 80, loss: 8.605829498264939e-05
step: 90, loss: 0.007154646795243025
step: 100, loss: 0.01997974142432213
step: 110, loss: 8.509570034220815e-05
step: 120, loss: 7.331358938245103e-05
step: 130, loss: 0.00010119762009708211
step: 140, loss: 0.008532600477337837
step: 150, loss: 6.153212598292157e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 160, loss: 5.798197162221186e-05
step: 170, loss: 0.00023958242672961205
/usr/local/lib/python3.7/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=0.837696335078534, f1=0.8753056234718827,

```

```
best_f1=0.8695652173913043
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.")
1916it [00:09, 209.61it/s]
load_f1 = 0.5584415584415584
real_f1 = 0.5102639296187683
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:31, 138.69it/s]
```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

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

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5891004800796509
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4428415894508362
step: 20, loss: 0.4634675681591034
step: 30, loss: 0.29766735434532166
step: 40, loss: 0.23626431822776794
step: 50, loss: 0.18549221754074097
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 60, loss: 0.13551673293113708
step: 70, loss: 0.20551268756389618
step: 80, loss: 0.029445283114910126
step: 90, loss: 0.07949160039424896
step: 100, loss: 0.2006663829088211
step: 110, loss: 0.10148154944181442
step: 120, loss: 0.009249869734048843
step: 130, loss: 0.1814500093460083
step: 140, loss: 0.028119364753365517
step: 150, loss: 0.14013755321502686
step: 160, loss: 0.06851822137832642
step: 170, loss: 0.09632544964551926
step: 180, loss: 0.045067332684993744
step: 190, loss: 0.007620927412062883
step: 200, loss: 0.027575917541980743
step: 210, loss: 0.039289116859436035
step: 220, loss: 0.10101057589054108
step: 230, loss: 0.006602437235414982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9620535714285715, f1=0.9641255605381166,
best_f1=0.9641255605381166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0036253745201975107
step: 10, loss: 0.1983962506055832
step: 20, loss: 0.01037950161844492
step: 30, loss: 0.00415817042812705
step: 40, loss: 0.06492824852466583
step: 50, loss: 0.014919349923729897
step: 60, loss: 0.026814274489879608
step: 70, loss: 0.05164596810936928
step: 80, loss: 0.008856951259076595
step: 90, loss: 0.06641816347837448
step: 100, loss: 0.017478352412581444
step: 110, loss: 0.02337072417140007
step: 120, loss: 0.026949914172291756
step: 130, loss: 0.02513563632965088
step: 140, loss: 0.004224635660648346
step: 150, loss: 0.07606562227010727
step: 160, loss: 0.01325895544141531
step: 170, loss: 0.004440372809767723
step: 180, loss: 0.002453444991260767
step: 190, loss: 0.0009248614078387618
step: 200, loss: 0.03147772327065468
step: 210, loss: 0.02760155312716961
step: 220, loss: 0.0029916574712842703
step: 230, loss: 0.0015887339832261205
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9668141592920354, f1=0.9686800894854586,
best_f1=0.9686800894854586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035498328506946564
step: 10, loss: 0.0032076395582407713
step: 20, loss: 0.007234751712530851
step: 30, loss: 0.00674245273694396
step: 40, loss: 0.06870704889297485
step: 50, loss: 0.008735097013413906
step: 60, loss: 0.017099829390645027
step: 70, loss: 0.017080791294574738
step: 80, loss: 0.1486699879169464
step: 90, loss: 0.07265950739383698
step: 100, loss: 0.006721820216625929
step: 110, loss: 0.07573940604925156

```

```

step: 120, loss: 0.06219755858182907
step: 130, loss: 0.010115080513060093
step: 140, loss: 0.0036489630583673716
step: 150, loss: 0.15978164970874786
step: 160, loss: 0.010922531597316265
step: 170, loss: 0.0017902874387800694
step: 180, loss: 0.008422212675213814
step: 190, loss: 0.006630973424762487
step: 200, loss: 0.012403870932757854
step: 210, loss: 0.008531699888408184
step: 220, loss: 0.12689681351184845
step: 230, loss: 0.009828914888203144
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9744160177975528, f1=0.9597315436241611,
best_f1=0.9597315436241611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0057143461890518665
step: 10, loss: 0.0010646753944456577
step: 20, loss: 0.0016974890604615211
step: 30, loss: 0.0014588473131880164
step: 40, loss: 0.02156807854771614
step: 50, loss: 0.015929076820611954
step: 60, loss: 0.009189709089696407
step: 70, loss: 0.05522925779223442
step: 80, loss: 0.13632002472877502
step: 90, loss: 0.09825877845287323
step: 100, loss: 0.020606568083167076
step: 110, loss: 0.0012541520409286022
step: 120, loss: 0.007805908564478159
step: 130, loss: 0.0024078215938061476
step: 140, loss: 0.0015886551700532436
step: 150, loss: 0.002876848680898547
step: 160, loss: 0.017413249239325523
step: 170, loss: 0.0010747454361990094
step: 180, loss: 0.11092819273471832
step: 190, loss: 0.004078340716660023
step: 200, loss: 0.1314510554075241
step: 210, loss: 0.0035671719815582037
step: 220, loss: 0.0021524287294596434
step: 230, loss: 0.007914496585726738

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9761634506242906, f1=0.952819332566168,
best_f1=0.952819332566168
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010829759994521737
step: 10, loss: 0.004440495278686285
step: 20, loss: 0.029126213863492012
step: 30, loss: 0.0019506918033584952
step: 40, loss: 0.01193302683532238
step: 50, loss: 0.0008438787772320211
step: 60, loss: 0.0031580915674567223
step: 70, loss: 0.003522673388943076
step: 80, loss: 0.01807571016252041
step: 90, loss: 0.015798307955265045
step: 100, loss: 0.000808258424513042
step: 110, loss: 0.007214678451418877
step: 120, loss: 0.0036555312108248472
step: 130, loss: 0.0005219049635343254
step: 140, loss: 0.0027873183134943247
step: 150, loss: 0.01793934777379036
step: 160, loss: 0.0017516508232802153
step: 170, loss: 0.0010052468860521913
step: 180, loss: 0.003512146882712841
step: 190, loss: 0.20424439013004303
step: 200, loss: 0.006665156688541174
step: 210, loss: 0.01335895899683237
step: 220, loss: 0.005130135454237461
step: 230, loss: 0.028948096558451653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776785714285714, f1=0.976324689966178,
best_f1=0.976324689966178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.0015278315404430032
step: 10, loss: 0.0023248200304806232
step: 20, loss: 0.0006222590454854071
step: 30, loss: 0.00030007632449269295
step: 40, loss: 0.00020919748931191862
step: 50, loss: 0.0002578684943728149
step: 60, loss: 0.0014273461420089006
step: 70, loss: 0.1589675396680832
step: 80, loss: 0.0020750188268721104
step: 90, loss: 0.0017382418736815453
step: 100, loss: 0.0005998389096930623
step: 110, loss: 0.004022924695163965
step: 120, loss: 0.00038224601303227246
step: 130, loss: 0.008364596404135227
step: 140, loss: 0.0031908873934298754
step: 150, loss: 0.0001671111094765365
step: 160, loss: 0.005482666660100222
step: 170, loss: 0.0010069309500977397
step: 180, loss: 0.004411119967699051
step: 190, loss: 0.0064531066454946995
step: 200, loss: 0.015462805517017841
step: 210, loss: 0.001850625965744257
step: 220, loss: 0.0033054444938898087
step: 230, loss: 0.0008609648793935776
/usr/local/lib/python3.7/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=0.9741863075196409, f1=0.9718785151856018,
best_f1=0.976324689966178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004914556629955769
step: 10, loss: 0.0005320276250131428
step: 20, loss: 0.00023416461772285402
step: 30, loss: 0.00030409314786083996
step: 40, loss: 0.04811355471611023
step: 50, loss: 0.0007355944253504276
step: 60, loss: 0.0007041132193990052
step: 70, loss: 0.003106695832684636
step: 80, loss: 0.0008642858592793345
step: 90, loss: 0.00040511126280762255
step: 100, loss: 0.0006495658308267593
step: 110, loss: 0.001138328923843801
step: 120, loss: 0.022057119756937027
step: 130, loss: 0.0014097915263846517
step: 140, loss: 0.0007362765609286726
step: 150, loss: 0.058781642466783524

```

```

step: 160, loss: 0.0003337115340400487
step: 170, loss: 0.025798119604587555
step: 180, loss: 0.0003010093350894749
step: 190, loss: 0.0007631039479747415
step: 200, loss: 0.006116360425949097
step: 210, loss: 0.0003821042482741177
step: 220, loss: 0.00048808014253154397
step: 230, loss: 0.0013953268062323332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.971815107102593,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014882689574733377
step: 10, loss: 0.0018609149847179651
step: 20, loss: 0.001990771619603038
step: 30, loss: 0.0015951694222167134
step: 40, loss: 0.0011442156974226236
step: 50, loss: 0.0006412086077034473
step: 60, loss: 0.0006997503223828971
step: 70, loss: 0.0003310233587399125
step: 80, loss: 0.00044280063593760133
step: 90, loss: 0.0003780114057008177
step: 100, loss: 0.0008858554065227509
step: 110, loss: 0.0003886317426804453
step: 120, loss: 0.00035839667543768883
step: 130, loss: 0.01156017929315567
step: 140, loss: 0.00020445416157599539
step: 150, loss: 0.056388191878795624
step: 160, loss: 0.0002156477712560445
step: 170, loss: 0.01623695157468319
step: 180, loss: 0.0001538364012958482
step: 190, loss: 0.0001575517817400396
step: 200, loss: 0.03159397095441818
step: 210, loss: 0.0007717291591688991
step: 220, loss: 0.0007326498162001371
step: 230, loss: 0.0009151221602223814
/usr/local/lib/python3.7/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=0.9766925638179801, f1=0.9675977653631285,

```

```

best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038278046995401382
step: 10, loss: 0.0064010475762188435
step: 20, loss: 0.0032481495290994644
step: 30, loss: 0.0021557325962930918
step: 40, loss: 0.009794346988201141
step: 50, loss: 0.002338527701795101
step: 60, loss: 0.0010956758633255959
step: 70, loss: 0.013362137600779533
step: 80, loss: 0.0016519823111593723
step: 90, loss: 0.04379939287900925
step: 100, loss: 0.0006704549305140972
step: 110, loss: 0.0005591055378317833
step: 120, loss: 0.03551703318953514
step: 130, loss: 0.028837701305747032
step: 140, loss: 0.0012275754706934094
step: 150, loss: 0.0014038683148100972
step: 160, loss: 0.0007273431983776391
step: 170, loss: 0.0001976129860850051
step: 180, loss: 0.002684136386960745
step: 190, loss: 0.000644797517452389
step: 200, loss: 0.0002057481906376779
step: 210, loss: 0.025904683396220207
step: 220, loss: 0.00020837862393818796
step: 230, loss: 9.978960588341579e-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=0.9787709497206705, f1=0.9751131221719457,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017394627502653748
step: 10, loss: 0.00025687809102237225
step: 20, loss: 0.0001778873265720904
step: 30, loss: 8.97091013030149e-05
step: 40, loss: 0.00028913840651512146
step: 50, loss: 6.486084748758003e-05
step: 60, loss: 0.00039192254189401865
step: 70, loss: 0.004916551988571882
step: 80, loss: 0.0007155667990446091
step: 90, loss: 0.00046688527800142765
step: 100, loss: 0.00036978020216338336
step: 110, loss: 0.00035555599606595933

```

```

step: 120, loss: 0.0003048541839234531
step: 130, loss: 0.0004438565520104021
step: 140, loss: 0.0037969325203448534
step: 150, loss: 0.0008670431561768055
step: 160, loss: 0.00014421960804611444
step: 170, loss: 0.0002488672616891563
step: 180, loss: 0.004291073884814978
step: 190, loss: 0.0004972670576535165
step: 200, loss: 0.0007240180275402963
step: 210, loss: 0.005784698761999607
step: 220, loss: 0.003027047263458371
step: 230, loss: 0.00024255330208688974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, f1=0.9739524348810873,
best_f1=0.9739524348810873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002391464076936245
step: 10, loss: 0.0018384744180366397
step: 20, loss: 0.000451504485681653
step: 30, loss: 0.000374637427739799
step: 40, loss: 0.00016023332136683166
step: 50, loss: 0.0009864246239885688
step: 60, loss: 0.0006355256773531437
step: 70, loss: 0.0005152865196578205
step: 80, loss: 0.00029110207105986774
step: 90, loss: 0.20147472620010376
step: 100, loss: 0.000587838760111481
step: 110, loss: 0.002618216909468174
step: 120, loss: 0.0005651566316373646
step: 130, loss: 0.0006341198459267616
step: 140, loss: 0.0003781883278861642
step: 150, loss: 0.000592827913351357
step: 160, loss: 0.0004492932348512113
step: 170, loss: 0.003424809779971838
step: 180, loss: 0.00032383896177634597
step: 190, loss: 0.00027041995781473815
step: 200, loss: 0.00035056297201663256
step: 210, loss: 0.000403938174713403
step: 220, loss: 0.0012928287032991648
step: 230, loss: 0.00026508228620514274

```

```

/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9718785151856018,
best_f1=0.9739524348810873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005542777944356203
step: 10, loss: 0.0002912183990702033
step: 20, loss: 0.0009708424331620336
step: 30, loss: 0.017302537336945534
step: 40, loss: 0.00037916662404313684
step: 50, loss: 0.0034838425926864147
step: 60, loss: 0.0003872636298183352
step: 70, loss: 0.0003271688474342227
step: 80, loss: 0.00011766517854994163
step: 90, loss: 0.0003388210607226938
step: 100, loss: 0.0003573241410776973
step: 110, loss: 0.00016035832231864333
step: 120, loss: 0.00028134521562606096
step: 130, loss: 0.00041629947372712195
step: 140, loss: 0.000864670320879668
step: 150, loss: 0.0008827264537103474
step: 160, loss: 0.0005900943651795387
step: 170, loss: 0.0006262261886149645
step: 180, loss: 0.0004937235498800874
step: 190, loss: 0.0009491863893344998
step: 200, loss: 0.0004135567578487098
step: 210, loss: 0.0024800398387014866
step: 220, loss: 0.02492474764585495
step: 230, loss: 0.0004470252606552094
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9821029082774049, f1=0.9740698985343857,
best_f1=0.9740698985343857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002795990149024874
step: 10, loss: 0.00039271486457437277
step: 20, loss: 0.0006421508733183146
step: 30, loss: 0.003463381202891469

```

```

step: 40, loss: 0.0010331105440855026
step: 50, loss: 0.0006242779199965298
step: 60, loss: 0.0002728290273807943
step: 70, loss: 0.0013686454622074962
step: 80, loss: 0.0001632019120734185
step: 90, loss: 0.00022543210070580244
step: 100, loss: 0.0005174529505893588
step: 110, loss: 0.0004980266094207764
step: 120, loss: 0.0004713325761258602
step: 130, loss: 0.0005290727131068707
step: 140, loss: 0.006490133237093687
step: 150, loss: 0.0003347054880578071
step: 160, loss: 0.0016416130820289254
step: 170, loss: 0.0012405901215970516
step: 180, loss: 0.019520243629813194
step: 190, loss: 0.0009871205547824502
step: 200, loss: 0.00026640217402018607
step: 210, loss: 0.0013244390720501542
step: 220, loss: 0.0006554904393851757
step: 230, loss: 0.0015845817979425192
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9751693002257337,
best_f1=0.9740698985343857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003735584323294461
step: 10, loss: 0.00038998006493784487
step: 20, loss: 0.00048559822607785463
step: 30, loss: 0.0007600918179377913
step: 40, loss: 0.00048725210945121944
step: 50, loss: 0.0004121062229387462
step: 60, loss: 0.0009643083321861923
step: 70, loss: 0.00040500378236174583
step: 80, loss: 0.00046541340998373926
step: 90, loss: 0.0008937673992477357
step: 100, loss: 0.0007659229449927807
step: 110, loss: 0.000877357495483011
step: 120, loss: 0.00017985753947868943
step: 130, loss: 0.0006497909780591726
step: 140, loss: 0.020691605284810066
step: 150, loss: 0.00022142514353618026
step: 160, loss: 0.0005989425117149949
step: 170, loss: 0.0006487545906566083
step: 180, loss: 0.0006002414156682789
step: 190, loss: 0.0004430457775015384

```

```

step: 200, loss: 0.0005002530524507165
step: 210, loss: 0.00039349074359051883
step: 220, loss: 0.0004744190664496273
step: 230, loss: 0.00022912293206900358
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9740112994350283,
best_f1=0.9740698985343857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005207396461628377
step: 10, loss: 0.00047161741531454027
step: 20, loss: 0.000727832259144634
step: 30, loss: 0.0005859676748514175
step: 40, loss: 0.00041509445873089135
step: 50, loss: 0.0005123650771565735
step: 60, loss: 0.004979949910193682
step: 70, loss: 0.0004310199583414942
step: 80, loss: 0.0008844878757372499
step: 90, loss: 0.00036414575879462063
step: 100, loss: 0.00028064247453585267
step: 110, loss: 0.0005481082480400801
step: 120, loss: 0.0071495939046144485
step: 130, loss: 0.00035669689532369375
step: 140, loss: 0.0007030823617242277
step: 150, loss: 0.0005212736432440579
step: 160, loss: 0.0006729582091793418
step: 170, loss: 0.00020355763263069093
step: 180, loss: 0.00057913304772228
step: 190, loss: 0.00022348760103341192
step: 200, loss: 0.0005450780736282468
step: 210, loss: 0.0052553629502654076
step: 220, loss: 0.0005520251579582691
step: 230, loss: 0.000438844901509583
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9740112994350283,
best_f1=0.9740698985343857
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

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

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:17, 137.66it/s]
load_f1 = 0.978865406006674
real_f1 = 0.9777777777777777
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 135.70it/s]

```

3.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```



```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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: 0.6491526365280151

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

step: 10, loss: 0.42966052889823914

step: 20, loss: 0.3214341402053833

step: 30, loss: 0.3791128695011139

step: 40, loss: 0.391519695520401

step: 50, loss: 0.6761741042137146

step: 60, loss: 0.3636443018913269

step: 70, loss: 0.4337458908557892

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

step: 80, loss: 0.16339415311813354

step: 90, loss: 0.21770863234996796

step: 100, loss: 0.3250778615474701

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

step: 110, loss: 0.19720633327960968

step: 120, loss: 0.2047078013420105

step: 130, loss: 0.14355477690696716

step: 140, loss: 0.3585885763168335

step: 150, loss: 0.14302562177181244

step: 160, loss: 0.1795421838760376

step: 170, loss: 0.09944013506174088

step: 180, loss: 0.07489364594221115

step: 190, loss: 0.10161466151475906

step: 200, loss: 0.07330726087093353

step: 210, loss: 0.07521487772464752

step: 220, loss: 0.177666574716568

step: 230, loss: 0.22774143517017365

step: 240, loss: 0.09721499681472778

step: 250, loss: 0.0381530299782753

step: 260, loss: 0.17739206552505493

step: 270, loss: 0.3305884599685669

step: 280, loss: 0.02893231064081192

step: 290, loss: 0.07464081048965454

step: 300, loss: 0.13556331396102905

step: 310, loss: 0.1924464851617813

step: 320, loss: 0.04541882872581482

step: 330, loss: 0.0833265408873558

step: 340, loss: 0.46184253692626953

step: 350, loss: 0.10589896142482758

step: 360, loss: 0.14137281477451324

```

step: 370, loss: 0.17565569281578064
step: 380, loss: 0.3104109764099121
step: 390, loss: 0.05809856206178665
step: 400, loss: 0.06762493401765823
step: 410, loss: 0.23038724064826965
step: 420, loss: 0.04140307754278183
step: 430, loss: 0.043083783239126205
step: 440, loss: 0.06509575247764587
step: 450, loss: 0.02784605883061886
step: 460, loss: 0.021357461810112
step: 470, loss: 0.026975462213158607
step: 480, loss: 0.09120529145002365
step: 490, loss: 0.047087159007787704
step: 500, loss: 0.03427926450967789
step: 510, loss: 0.044501882046461105
step: 520, loss: 0.02368675172328949
step: 530, loss: 0.05954144150018692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9374130737134909, f1=0.9385474860335196,
best_f1=0.9385474860335196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04833117872476578
step: 10, loss: 0.08893131464719772
step: 20, loss: 0.0965239554643631
step: 30, loss: 0.09201759845018387
step: 40, loss: 0.06548429280519485
step: 50, loss: 0.06835055351257324
step: 60, loss: 0.03111908584833145
step: 70, loss: 0.030984481796622276
step: 80, loss: 0.025081077590584755
step: 90, loss: 0.036526866257190704
step: 100, loss: 0.16097721457481384
step: 110, loss: 0.005680826958268881
step: 120, loss: 0.013409789651632309
step: 130, loss: 0.0021936355624347925
step: 140, loss: 0.15815596282482147
step: 150, loss: 0.03200111538171768
step: 160, loss: 0.026012741029262543
step: 170, loss: 0.03878254070878029
step: 180, loss: 0.05229726433753967

```

```

step: 190, loss: 0.01722356677055359
step: 200, loss: 0.3083826005458832
step: 210, loss: 0.013074413873255253
step: 220, loss: 0.0037677015643566847
step: 230, loss: 0.013552680611610413
step: 240, loss: 0.005519213620573282
step: 250, loss: 0.015621678903698921
step: 260, loss: 0.040193893015384674
step: 270, loss: 0.012779662385582924
step: 280, loss: 0.15092061460018158
step: 290, loss: 0.0424979142844677
step: 300, loss: 0.0768866166472435
step: 310, loss: 0.022220442071557045
step: 320, loss: 0.07091069221496582
step: 330, loss: 0.06606373935937881
step: 340, loss: 0.04860867187380791
step: 350, loss: 0.00299275666475296
step: 360, loss: 0.09346690773963928
step: 370, loss: 0.01612769067287445
step: 380, loss: 0.08748997002840042
step: 390, loss: 0.003104636212810874
step: 400, loss: 0.11935313791036606
step: 410, loss: 0.11990699172019958
step: 420, loss: 0.03794903680682182
step: 430, loss: 0.165157288312912
step: 440, loss: 0.0034976042807102203
step: 450, loss: 0.04836495220661163
step: 460, loss: 0.052996233105659485
step: 470, loss: 0.04631591960787773
step: 480, loss: 0.011053401976823807
step: 490, loss: 0.07697469741106033
step: 500, loss: 0.0060097835958004
step: 510, loss: 0.008738423697650433
step: 520, loss: 0.15912163257598877
step: 530, loss: 0.037082646042108536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9462068965517241, f1=0.9474169741697417,
best_f1=0.9474169741697417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1250765174627304

```

step: 10, loss: 0.015103642828762531
step: 20, loss: 0.04054257646203041
step: 30, loss: 0.08030987530946732
step: 40, loss: 0.08571026474237442
step: 50, loss: 0.002782852854579687
step: 60, loss: 0.048279885202646255
step: 70, loss: 0.002722123172134161
step: 80, loss: 0.09990181028842926
step: 90, loss: 0.01770283840596676
step: 100, loss: 0.023838605731725693
step: 110, loss: 0.028103096410632133
step: 120, loss: 0.1749795377254486
step: 130, loss: 0.009374367073178291
step: 140, loss: 0.03429070860147476
step: 150, loss: 0.023538200184702873
step: 160, loss: 0.012811439111828804
step: 170, loss: 0.012396365404129028
step: 180, loss: 0.07055115699768066
step: 190, loss: 0.013735630549490452
step: 200, loss: 0.044784024357795715
step: 210, loss: 0.015353845432400703
step: 220, loss: 0.12636099755764008
step: 230, loss: 0.06811391562223434
step: 240, loss: 0.12665686011314392
step: 250, loss: 0.20215845108032227
step: 260, loss: 0.13492009043693542
step: 270, loss: 0.00379091571085155
step: 280, loss: 0.06948499381542206
step: 290, loss: 0.013705579563975334
step: 300, loss: 0.17852665483951569
step: 310, loss: 0.03993232920765877
step: 320, loss: 0.010807141661643982
step: 330, loss: 0.010357380844652653
step: 340, loss: 0.009326335974037647
step: 350, loss: 0.072162926197052
step: 360, loss: 0.02093052677810192
step: 370, loss: 0.0713321790099144
step: 380, loss: 0.005146999843418598
step: 390, loss: 0.005994261708110571
step: 400, loss: 0.07489193975925446
step: 410, loss: 0.05083842948079109
step: 420, loss: 0.006335200741887093
step: 430, loss: 0.011989270336925983
step: 440, loss: 0.17367862164974213
step: 450, loss: 0.05099098011851311
step: 460, loss: 0.11657581478357315
step: 470, loss: 0.03079918399453163
step: 480, loss: 0.22766394913196564

```

step: 490, loss: 0.06787559390068054
step: 500, loss: 0.00808112695813179
step: 510, loss: 0.05393407121300697
step: 520, loss: 0.0034049360547214746
step: 530, loss: 0.008155963383615017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.948087431693989, f1=0.942035600182565,
best_f1=0.942035600182565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017097918316721916
step: 10, loss: 0.0057921395637094975
step: 20, loss: 0.019388379529118538
step: 30, loss: 0.12726058065891266
step: 40, loss: 0.02220282517373562
step: 50, loss: 0.03957900404930115
step: 60, loss: 0.030942648649215698
step: 70, loss: 0.22264394164085388
step: 80, loss: 0.11005955934524536
step: 90, loss: 0.09803146123886108
step: 100, loss: 0.002306475769728422
step: 110, loss: 0.03265872225165367
step: 120, loss: 0.01208868995308876
step: 130, loss: 0.08096424490213394
step: 140, loss: 0.06405545771121979
step: 150, loss: 0.03185221552848816
step: 160, loss: 0.030962683260440826
step: 170, loss: 0.008163665421307087
step: 180, loss: 0.11263397336006165
step: 190, loss: 0.04678668454289436
step: 200, loss: 0.09676321595907211
step: 210, loss: 0.00299422862008214
step: 220, loss: 0.04842314496636391
step: 230, loss: 0.008686280809342861
step: 240, loss: 0.03344310447573662
step: 250, loss: 0.006909291725605726
step: 260, loss: 0.0010827162768691778
step: 270, loss: 0.1785447597503662
step: 280, loss: 0.010468407534062862
step: 290, loss: 0.0437772199511528
step: 300, loss: 0.0069020469672977924

```

```

step: 310, loss: 0.00690125860273838
step: 320, loss: 0.028349444270133972
step: 330, loss: 0.06918660551309586
step: 340, loss: 0.0010703192092478275
step: 350, loss: 0.1317494511604309
step: 360, loss: 0.1184052899479866
step: 370, loss: 0.0032555253710597754
step: 380, loss: 0.002820805413648486
step: 390, loss: 0.0016983187524601817
step: 400, loss: 0.09812553226947784
step: 410, loss: 0.002067424124106765
step: 420, loss: 0.06235959380865097
step: 430, loss: 0.04636798053979874
step: 440, loss: 0.0030799084343016148
step: 450, loss: 0.011483849957585335
step: 460, loss: 0.009450550191104412
step: 470, loss: 0.001526394160464406
step: 480, loss: 0.01082915160804987
step: 490, loss: 0.0056231156922876835
step: 500, loss: 0.03892508149147034
step: 510, loss: 0.0049062613397836685
step: 520, loss: 0.019762463867664337
step: 530, loss: 0.14126752316951752
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9596622889305815, f1=0.948356807511737,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006484378129243851
step: 10, loss: 0.04237828403711319
step: 20, loss: 0.0012448757188394666
step: 30, loss: 0.014078404754400253
step: 40, loss: 0.0008059407118707895
step: 50, loss: 0.06486999988555908
step: 60, loss: 0.02563145011663437
step: 70, loss: 0.016178887337446213
step: 80, loss: 0.0027675852179527283
step: 90, loss: 0.020186936482787132
step: 100, loss: 0.01675577647984028
step: 110, loss: 0.008182721212506294
step: 120, loss: 0.15617358684539795

```

```

step: 130, loss: 0.021025095134973526
step: 140, loss: 0.02759459987282753
step: 150, loss: 0.008743987418711185
step: 160, loss: 0.0028504091314971447
step: 170, loss: 0.011054172180593014
step: 180, loss: 0.010235331952571869
step: 190, loss: 0.007497337181121111
step: 200, loss: 0.0029635014943778515
step: 210, loss: 0.0018517090938985348
step: 220, loss: 0.0070450324565172195
step: 230, loss: 0.0013160653179511428
step: 240, loss: 0.006409573368728161
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 250, loss: 0.3246553838253021
step: 260, loss: 0.0027585134375840425
step: 270, loss: 0.0021660609636455774
step: 280, loss: 0.016901927068829536
step: 290, loss: 0.09018855541944504
step: 300, loss: 0.02966323122382164
step: 310, loss: 0.012882005423307419
step: 320, loss: 0.004612443037331104
step: 330, loss: 0.00046100205508992076
step: 340, loss: 0.04698082432150841
step: 350, loss: 0.004229849670082331
step: 360, loss: 0.0010443764040246606
step: 370, loss: 0.0023564640432596207
step: 380, loss: 0.0008454180206172168
step: 390, loss: 0.0044301762245595455
step: 400, loss: 0.005133284255862236
step: 410, loss: 0.047880083322525024
step: 420, loss: 0.1688830852508545
step: 430, loss: 0.004550556186586618
step: 440, loss: 0.03607596084475517
step: 450, loss: 0.004805159289389849
step: 460, loss: 0.10423818975687027
step: 470, loss: 0.06346560269594193
step: 480, loss: 0.023746749386191368
step: 490, loss: 0.018413856625556946
step: 500, loss: 0.0007377452566288412
step: 510, loss: 0.0037185123655945063
step: 520, loss: 0.27414411306381226
step: 530, loss: 0.01746106892824173
/usr/local/lib/python3.7/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=0.9466357308584686, f1=0.9436749769159741,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```


An input tensor was not cuda.

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

```
step: 0, loss: 0.05815211310982704
step: 10, loss: 0.0036940020509064198
step: 20, loss: 0.0030863790307193995
step: 30, loss: 0.0007204034482128918
step: 40, loss: 0.03243004158139229
step: 50, loss: 0.0007259981357492507
step: 60, loss: 0.0019685719162225723
step: 70, loss: 0.0005169707001186907
step: 80, loss: 0.0004246525058988482
step: 90, loss: 0.0024709301069378853
step: 100, loss: 0.0010374098783358932
step: 110, loss: 0.0009109490201808512
step: 120, loss: 0.022070197388529778
step: 130, loss: 0.012900453992187977
step: 140, loss: 0.004350211005657911
step: 150, loss: 0.028986921533942223
step: 160, loss: 0.04474835470318794
step: 170, loss: 0.004844271577894688
step: 180, loss: 0.010441748425364494
step: 190, loss: 0.19190320372581482
step: 200, loss: 0.0035069396253675222
step: 210, loss: 0.004714600276201963
step: 220, loss: 0.00784775149077177
step: 230, loss: 0.0011061258846893907
step: 240, loss: 0.011653969995677471
step: 250, loss: 0.0020998925901949406
step: 260, loss: 0.0074006798677146435
step: 270, loss: 0.0008902983390726149
step: 280, loss: 0.0016694720834493637
step: 290, loss: 0.010291142389178276
step: 300, loss: 0.02944730781018734
step: 310, loss: 0.05608050897717476
step: 320, loss: 0.0002582386659923941
step: 330, loss: 0.00042752944864332676
step: 340, loss: 0.0001282142475247383
step: 350, loss: 0.003119983710348606
step: 360, loss: 0.0007069393759593368
step: 370, loss: 0.023803353309631348
step: 380, loss: 0.005015337374061346
step: 390, loss: 0.01849723607301712
step: 400, loss: 0.022474156692624092
step: 410, loss: 0.0008675775025039911
step: 420, loss: 0.000890958821401
step: 430, loss: 0.041735660284757614
step: 440, loss: 0.002032829448580742
step: 450, loss: 0.17954561114311218
```

```
step: 460, loss: 0.004099463578313589
step: 470, loss: 0.0027164986822754145
step: 480, loss: 0.005386188626289368
step: 490, loss: 0.006072087679058313
step: 500, loss: 0.014838552102446556
step: 510, loss: 0.1508990377187729
step: 520, loss: 0.0032240785658359528
step: 530, loss: 0.004359085112810135
/usr/local/lib/python3.7/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=0.9478584729981377, f1=0.9376163873370577,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021644635125994682
step: 10, loss: 0.0008458970696665347
step: 20, loss: 0.009444255381822586
step: 30, loss: 0.003038086462765932
step: 40, loss: 0.0057996404357254505
step: 50, loss: 0.03916357085108757
step: 60, loss: 0.0025465500075370073
step: 70, loss: 0.011877084150910378
step: 80, loss: 0.0008351026917807758
step: 90, loss: 0.0013574989279732108
step: 100, loss: 0.0030996461864560843
step: 110, loss: 0.0004198820679448545
step: 120, loss: 0.002221280476078391
step: 130, loss: 0.0003845289465971291
step: 140, loss: 0.0008117405814118683
step: 150, loss: 0.049897074699401855
step: 160, loss: 0.0002940080303233117
step: 170, loss: 0.009934872388839722
step: 180, loss: 0.13220283389091492
step: 190, loss: 0.01328936405479908
step: 200, loss: 0.00018060543516185135
step: 210, loss: 0.06819842755794525
step: 220, loss: 0.0001379907043883577
step: 230, loss: 0.0070346649736166
step: 240, loss: 0.017064841464161873
step: 250, loss: 0.13005836308002472
step: 260, loss: 0.026841850951313972
step: 270, loss: 0.0025652358308434486
step: 280, loss: 0.00563808111473918
step: 290, loss: 0.002940954640507698
step: 300, loss: 0.0017788023687899113
step: 310, loss: 0.0010103201493620872
```

```

step: 320, loss: 0.0031299942638725042
step: 330, loss: 0.0006117494776844978
step: 340, loss: 0.0002457398804835975
step: 350, loss: 0.0003928454243578017
step: 360, loss: 0.0006733086192980409
step: 370, loss: 0.0004154439375270158
step: 380, loss: 0.011831841431558132
step: 390, loss: 0.017211206257343292
step: 400, loss: 0.029254641383886337
step: 410, loss: 0.0019440613687038422
step: 420, loss: 0.09341221302747726
step: 430, loss: 0.007998713292181492
step: 440, loss: 0.013896690681576729
step: 450, loss: 0.012271150015294552
step: 460, loss: 0.001511072157882154
step: 470, loss: 0.1382066160440445
step: 480, loss: 0.0016025914810597897
step: 490, loss: 0.010583990253508091
step: 500, loss: 0.0008356185280717909
step: 510, loss: 0.0013795720878988504
step: 520, loss: 0.004150224383920431
step: 530, loss: 0.0024057822301983833
/usr/local/lib/python3.7/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=0.9516279069767443, f1=0.9483960948396094,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001871276879683137
step: 10, loss: 0.0024253069423139095
step: 20, loss: 0.0016660174587741494
step: 30, loss: 0.0008100740960799158
step: 40, loss: 0.0003296906652394682
step: 50, loss: 0.00025563116651028395
step: 60, loss: 0.00043132115388289094
step: 70, loss: 0.0022406606003642082
step: 80, loss: 0.007075293455272913
step: 90, loss: 0.0002534120576456189
step: 100, loss: 0.0020379782654345036
step: 110, loss: 0.0006572067504748702
step: 120, loss: 0.0017595579847693443
step: 130, loss: 0.0003076008288189769
step: 140, loss: 0.015980619937181473
step: 150, loss: 0.09489081799983978
step: 160, loss: 0.00024872980429790914
step: 170, loss: 0.014491516165435314

```

```

step: 180, loss: 0.00034264501300640404
step: 190, loss: 0.006839381996542215
step: 200, loss: 0.0026626414619386196
step: 210, loss: 0.2508850693702698
step: 220, loss: 0.0033937639091163874
step: 230, loss: 0.055090438574552536
step: 240, loss: 0.0016477819299325347
step: 250, loss: 0.0018970266683027148
step: 260, loss: 0.0004539554938673973
step: 270, loss: 0.07263660430908203
step: 280, loss: 0.0015862053260207176
step: 290, loss: 0.0012471158988773823
step: 300, loss: 7.215363439172506e-05
step: 310, loss: 0.0002991874353028834
step: 320, loss: 0.00019750256615225226
step: 330, loss: 0.0006047209026291966
step: 340, loss: 0.042326439172029495
step: 350, loss: 9.352238703286275e-05
step: 360, loss: 0.017029568552970886
step: 370, loss: 0.03479698672890663
step: 380, loss: 6.325679714791477e-05
step: 390, loss: 0.0013989099534228444
step: 400, loss: 0.0021067315246909857
step: 410, loss: 0.014149002730846405
step: 420, loss: 0.001018937909975648
step: 430, loss: 0.027536818757653236
step: 440, loss: 0.0102397371083498
step: 450, loss: 0.001375120016746223
step: 460, loss: 0.0015729565639048815
step: 470, loss: 0.014106407761573792
step: 480, loss: 0.0024094569962471724
step: 490, loss: 0.08005446940660477
step: 500, loss: 0.001553592155687511
step: 510, loss: 0.002071579685434699
step: 520, loss: 0.0004050632705911994
step: 530, loss: 0.0013559508370235562
/usr/local/lib/python3.7/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=0.9541029207232267, f1=0.9478060046189377,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014590873615816236
step: 10, loss: 0.0013663733843713999
step: 20, loss: 0.00033509134664200246
step: 30, loss: 0.0925876572728157

```

step: 40, loss: 0.0005362362135201693
step: 50, loss: 0.0005320553900673985
step: 60, loss: 0.0016445884248241782
step: 70, loss: 0.05292157828807831
step: 80, loss: 0.06996151059865952
step: 90, loss: 0.015721336007118225
step: 100, loss: 0.019876981154084206
step: 110, loss: 0.003441284643486142
step: 120, loss: 0.0007430510013364255
step: 130, loss: 0.0003606991667766124
step: 140, loss: 0.0003971426631323993
step: 150, loss: 0.00042121735168620944
step: 160, loss: 0.13109421730041504
step: 170, loss: 0.022091900929808617
step: 180, loss: 0.005753659177571535
step: 190, loss: 0.00023140120902098715
step: 200, loss: 0.00040039524901658297
step: 210, loss: 0.00022884267673362046
step: 220, loss: 0.0011472837068140507
step: 230, loss: 0.00865130964666605
step: 240, loss: 0.00024852954084053636
step: 250, loss: 0.021723993122577667
step: 260, loss: 0.00020285767095629126
step: 270, loss: 0.004636791534721851
step: 280, loss: 0.001996305538341403
step: 290, loss: 0.00012760108802467585
step: 300, loss: 0.0017894773045554757
step: 310, loss: 0.20444045960903168
step: 320, loss: 0.00996061135083437
step: 330, loss: 0.00046684587141498923
step: 340, loss: 0.0003875379334203899
step: 350, loss: 0.004593688528984785
step: 360, loss: 0.0001120672604884021
step: 370, loss: 0.00010349402873544022
step: 380, loss: 0.000788002391345799
step: 390, loss: 0.00033233032445423305
step: 400, loss: 0.0013024525251239538
step: 410, loss: 0.0010223948629572988
step: 420, loss: 0.00011796918988693506
step: 430, loss: 0.00034753093495965004
step: 440, loss: 6.567218224518001e-05
step: 450, loss: 0.013313453644514084
step: 460, loss: 7.924372766865417e-05
step: 470, loss: 5.939418406342156e-05
step: 480, loss: 0.00012202688958495855
step: 490, loss: 0.046264518052339554
step: 500, loss: 0.00012499201693572104
step: 510, loss: 7.699980778852478e-05

```
step: 520, loss: 0.02785784751176834
step: 530, loss: 0.050598837435245514
/usr/local/lib/python3.7/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=0.9487771112136594, f1=0.9446730681298583,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005154603626579046
step: 10, loss: 0.002204491524025798
step: 20, loss: 0.003984004724770784
step: 30, loss: 0.002624194836243987
step: 40, loss: 0.00027886719908565283
step: 50, loss: 0.00019851946854032576
step: 60, loss: 0.003185539273545146
step: 70, loss: 7.155588536988944e-05
step: 80, loss: 5.4848056606715545e-05
step: 90, loss: 0.0023072687909007072
step: 100, loss: 0.006977993994951248
step: 110, loss: 0.014193912036716938
step: 120, loss: 3.931088212993927e-05
step: 130, loss: 0.0015901810256764293
step: 140, loss: 0.0005083174328319728
step: 150, loss: 0.0018454118398949504
step: 160, loss: 0.004543599672615528
step: 170, loss: 0.00020561114070005715
step: 180, loss: 0.00019909856200683862
step: 190, loss: 0.00011478488158900291
step: 200, loss: 0.005447934381663799
step: 210, loss: 0.016171418130397797
step: 220, loss: 0.0014931376790627837
step: 230, loss: 0.0008402678067795932
step: 240, loss: 0.008943093940615654
step: 250, loss: 0.0005150712677277625
step: 260, loss: 0.001655848347581923
step: 270, loss: 0.008428161963820457
step: 280, loss: 0.0071356333792209625
step: 290, loss: 0.0007028339896351099
step: 300, loss: 0.004636804573237896
step: 310, loss: 0.004956393502652645
step: 320, loss: 0.0034769924823194742
step: 330, loss: 0.017812877893447876
step: 340, loss: 0.056476593017578125
step: 350, loss: 0.0016676264349371195
step: 360, loss: 8.68847273522988e-05
step: 370, loss: 0.002061012201011181
```

```

step: 380, loss: 0.001401590183377266
step: 390, loss: 7.164628186728805e-05
step: 400, loss: 0.00017973626381717622
step: 410, loss: 0.0003341316187288612
step: 420, loss: 0.015398350544273853
step: 430, loss: 0.0004567840078379959
step: 440, loss: 0.0012672626180574298
step: 450, loss: 0.00016489824338350445
step: 460, loss: 0.0006228192360140383
step: 470, loss: 0.0010450065601617098
step: 480, loss: 0.00029489968437701464
step: 490, loss: 0.0009639033814892173
step: 500, loss: 0.0001360874157398939
step: 510, loss: 6.601821951335296e-05
step: 520, loss: 0.00022017641458660364
step: 530, loss: 0.0008193730027414858
/usr/local/lib/python3.7/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=0.949456778460085, f1=0.9446284902981543,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016712395881768316
step: 10, loss: 0.001108388532884419
step: 20, loss: 6.169991684146225e-05
step: 30, loss: 0.003336873836815357
step: 40, loss: 7.890420965850353e-05
step: 50, loss: 4.8286980018019676e-05
step: 60, loss: 2.72929228231078e-05
step: 70, loss: 0.00014772408758290112
step: 80, loss: 0.0006570580299012363
step: 90, loss: 0.000689181499183178
step: 100, loss: 0.00035016078618355095
step: 110, loss: 6.298408698057756e-05
step: 120, loss: 0.0007636989466845989
step: 130, loss: 8.739429904380813e-05
step: 140, loss: 5.567731204791926e-05
step: 150, loss: 4.3596417526714504e-05
step: 160, loss: 0.0004071332514286041
step: 170, loss: 0.0018879021517932415
step: 180, loss: 0.00021247727272566408
step: 190, loss: 0.0003635639150161296
step: 200, loss: 7.716139953117818e-05
step: 210, loss: 8.95224220585078e-05
step: 220, loss: 0.059797532856464386
step: 230, loss: 9.809842595132068e-05

```

```

step: 240, loss: 0.0022227393928915262
step: 250, loss: 0.0006535633001476526
step: 260, loss: 0.00426107831299305
step: 270, loss: 0.002698234049603343
step: 280, loss: 0.0005045172874815762
step: 290, loss: 0.000450472958618775
step: 300, loss: 0.00043504475615918636
step: 310, loss: 0.2195265144109726
step: 320, loss: 0.0010880022309720516
step: 330, loss: 0.0002550424251239747
step: 340, loss: 0.0011817245976999402
step: 350, loss: 0.000675079645588994
step: 360, loss: 0.0005276550655253232
step: 370, loss: 0.0015056036645546556
step: 380, loss: 0.001693460624665022
step: 390, loss: 0.0012694016331806779
step: 400, loss: 0.0007918644696474075
step: 410, loss: 0.0011540402192622423
step: 420, loss: 0.001110742217861116
step: 430, loss: 0.0027958801947534084
step: 440, loss: 0.00031588852289132774
step: 450, loss: 0.001616118592210114
step: 460, loss: 0.0006350466283038259
step: 470, loss: 0.0009607558022253215
step: 480, loss: 0.0017280633328482509
step: 490, loss: 0.00020071477047167718
step: 500, loss: 0.0017837831983342767
step: 510, loss: 0.006115948315709829
step: 520, loss: 0.001378362183459103
step: 530, loss: 0.013242006301879883
/usr/local/lib/python3.7/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=0.9517625231910947, f1=0.9425393883225208,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024227635003626347
step: 10, loss: 0.005571101792156696
step: 20, loss: 0.001037565409205854
step: 30, loss: 7.208088936749846e-05
step: 40, loss: 0.0013393587432801723
step: 50, loss: 8.206696656998247e-05
step: 60, loss: 0.0001334300177404657
step: 70, loss: 0.0012220091884955764
step: 80, loss: 0.00012434207019396126
step: 90, loss: 0.00024938557180576026

```



```
step: 100, loss: 0.24834595620632172
step: 110, loss: 0.00015885330503806472
step: 120, loss: 0.002963718492537737
step: 130, loss: 0.0017172226216644049
step: 140, loss: 0.0015235078753903508
step: 150, loss: 0.00018893145897891372
step: 160, loss: 0.00321813952177763
step: 170, loss: 0.00012864005111623555
step: 180, loss: 0.0001250109780812636
step: 190, loss: 0.0037718256935477257
step: 200, loss: 0.0022034042049199343
step: 210, loss: 0.00013363553443923593
step: 220, loss: 0.00012392857752274722
step: 230, loss: 9.175492596114054e-05
step: 240, loss: 0.03035084716975689
step: 250, loss: 9.31762988329865e-05
step: 260, loss: 0.0007306287297978997
step: 270, loss: 0.0009789887117221951
step: 280, loss: 0.0003379075787961483
step: 290, loss: 0.0005028376472182572
step: 300, loss: 0.0038800423499196768
step: 310, loss: 0.0001849018008215353
step: 320, loss: 0.007379262242466211
step: 330, loss: 0.0018461920553818345
step: 340, loss: 0.002155248075723648
step: 350, loss: 0.000538715859875083
step: 360, loss: 0.0004671070200856775
step: 370, loss: 6.134924478828907e-05
step: 380, loss: 0.002244328847154975
step: 390, loss: 0.0003985915973316878
step: 400, loss: 7.652290514670312e-05
step: 410, loss: 0.0032637265976518393
step: 420, loss: 0.00907284114509821
step: 430, loss: 0.0004450421838555485
step: 440, loss: 0.0012329414021223783
step: 450, loss: 0.0002187084755860269
step: 460, loss: 0.00010212411871179938
step: 470, loss: 0.0015876634279266
step: 480, loss: 0.0001737951533868909
step: 490, loss: 8.915551734389737e-05
step: 500, loss: 5.059749310021289e-05
step: 510, loss: 0.0050760312005877495
step: 520, loss: 0.0034069251269102097
step: 530, loss: 0.004061132203787565
/usr/local/lib/python3.7/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=0.9539748953974895, f1=0.9434137291280148,
```

```
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.830186273669824e-05
step: 10, loss: 9.411084465682507e-05
step: 20, loss: 0.0004892387660220265
step: 30, loss: 6.0255901189520955e-05
step: 40, loss: 0.006460313219577074
step: 50, loss: 0.006922323722392321
step: 60, loss: 9.573811257723719e-05
step: 70, loss: 6.839606794528663e-05
step: 80, loss: 0.00043973367428407073
step: 90, loss: 0.0027384802233427763
step: 100, loss: 6.0726564697688445e-05
step: 110, loss: 4.2267209209967405e-05
step: 120, loss: 8.759919728618115e-05
step: 130, loss: 4.0928793168859556e-05
step: 140, loss: 7.679215195821598e-05
step: 150, loss: 0.0016473227879032493
step: 160, loss: 0.0002094705996569246
step: 170, loss: 0.0008506203885190189
step: 180, loss: 0.0002066524320980534
step: 190, loss: 0.00010913913138210773
step: 200, loss: 8.770901331445202e-05
step: 210, loss: 5.7128785556415096e-05
step: 220, loss: 5.355707980925217e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 0.0011349869892001152
step: 240, loss: 0.0008824800024740398
step: 250, loss: 0.0011625366751104593
step: 260, loss: 0.000490404199808836
step: 270, loss: 5.6228451285278425e-05
step: 280, loss: 5.114030864206143e-05
step: 290, loss: 7.762199675198644e-05
step: 300, loss: 5.413110557128675e-05
step: 310, loss: 6.257439235923812e-05
step: 320, loss: 6.170599954202771e-05
step: 330, loss: 6.205507816048339e-05
step: 340, loss: 0.003641868242993951
step: 350, loss: 4.0332164644496515e-05
step: 360, loss: 0.1254616379737854
step: 370, loss: 8.685739885549992e-05
step: 380, loss: 0.0005734653677791357
step: 390, loss: 0.00022973031445872039
step: 400, loss: 0.0014278801390901208
step: 410, loss: 0.0003534159332048148
step: 420, loss: 7.525015826104209e-05
```

```
step: 430, loss: 0.0002881059190258384
step: 440, loss: 8.998671546578407e-05
step: 450, loss: 0.00013205318828113377
step: 460, loss: 0.0037152059376239777
step: 470, loss: 0.001433345372788608
step: 480, loss: 4.1608971514506266e-05
step: 490, loss: 6.262358510866761e-05
step: 500, loss: 4.408536187838763e-05
step: 510, loss: 7.556454511359334e-05
step: 520, loss: 5.748072362621315e-05
step: 530, loss: 5.992351725581102e-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=0.9506641366223909, f1=0.9435215946843853,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009015500545501709
step: 10, loss: 0.0004867248353548348
step: 20, loss: 6.699713412672281e-05
step: 30, loss: 7.522809755755588e-05
step: 40, loss: 0.0005618127179332078
step: 50, loss: 0.00012290265294723213
step: 60, loss: 6.720302917528898e-05
step: 70, loss: 8.55637263157405e-05
step: 80, loss: 8.747982792556286e-05
step: 90, loss: 4.136948336963542e-05
step: 100, loss: 5.8451019867789e-05
step: 110, loss: 5.781775325885974e-05
step: 120, loss: 6.678348290733993e-05
step: 130, loss: 0.00030932543450035155
step: 140, loss: 8.712247654329985e-05
step: 150, loss: 0.0013287357287481427
step: 160, loss: 8.78383289091289e-05
step: 170, loss: 6.238237983779982e-05
step: 180, loss: 6.957594450796023e-05
step: 190, loss: 4.691056165029295e-05
step: 200, loss: 8.076983067439869e-05
step: 210, loss: 9.730157034937292e-05
step: 220, loss: 3.9540744182886556e-05
step: 230, loss: 4.8293837608071044e-05
step: 240, loss: 0.08499671518802643
step: 250, loss: 0.0010021600173786283
step: 260, loss: 0.0002411987807136029
step: 270, loss: 0.00033341324888169765
step: 280, loss: 6.534165004268289e-05
```

```

step: 290, loss: 5.082254210719839e-05
step: 300, loss: 0.00011373471352271736
step: 310, loss: 4.7437468310818076e-05
step: 320, loss: 0.00013565525296144187
step: 330, loss: 0.0021012118086218834
step: 340, loss: 0.0010242452844977379
step: 350, loss: 6.358598329825327e-05
step: 360, loss: 0.0013820755993947387
step: 370, loss: 0.0018293465254828334
step: 380, loss: 0.0011868159053847194
step: 390, loss: 0.00023364427033811808
step: 400, loss: 0.0038938079960644245
step: 410, loss: 4.388467641547322e-05
step: 420, loss: 6.220286741154268e-05
step: 430, loss: 0.0033587513025850058
step: 440, loss: 0.0013240452390164137
step: 450, loss: 0.00024875413510017097
step: 460, loss: 0.0027842307463288307
step: 470, loss: 9.31407994357869e-05
step: 480, loss: 0.00011733771680155769
step: 490, loss: 8.211256499635056e-05
step: 500, loss: 0.0021408561151474714
step: 510, loss: 0.0007657846435904503
step: 520, loss: 0.0016840153839439154
step: 530, loss: 0.0037918591406196356
/usr/local/lib/python3.7/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=0.954060324825986, f1=0.9499536607970344,
best_f1=0.948356807511737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001638360321521759
step: 10, loss: 0.00218755379319191
step: 20, loss: 0.002839829074218869
step: 30, loss: 0.009150831960141659
step: 40, loss: 8.263331255875528e-05
step: 50, loss: 0.0009573168936185539
step: 60, loss: 0.0005823491956107318
step: 70, loss: 6.69145883875899e-05
step: 80, loss: 5.749686897615902e-05
step: 90, loss: 0.0001323867036262527
step: 100, loss: 2.6245535991620272e-05
step: 110, loss: 0.0012305309064686298
step: 120, loss: 5.064615106675774e-05
step: 130, loss: 0.0003453497192822397
step: 140, loss: 4.9063914048019797e-05

```

```

step: 150, loss: 9.199138730764389e-05
step: 160, loss: 7.292097143363208e-05
step: 170, loss: 3.6994784750277176e-05
step: 180, loss: 7.710505451541394e-05
step: 190, loss: 0.0007931714644655585
step: 200, loss: 0.0011426301207393408
step: 210, loss: 0.00040785191231407225
step: 220, loss: 8.074868674157187e-05
step: 230, loss: 8.675752906128764e-05
step: 240, loss: 5.08676930621732e-05
step: 250, loss: 8.481126133119687e-05
step: 260, loss: 3.929094600607641e-05
step: 270, loss: 0.00796664971858263
step: 280, loss: 4.413661008584313e-05
step: 290, loss: 0.00018618570175021887
step: 300, loss: 6.949021189939231e-05
step: 310, loss: 0.054995160549879074
step: 320, loss: 6.246293196454644e-05
step: 330, loss: 7.908639963716269e-05
step: 340, loss: 0.00011789743439294398
step: 350, loss: 0.004937917459756136
step: 360, loss: 0.00018572768021840602
step: 370, loss: 0.00010256384848617017
step: 380, loss: 6.646917609032243e-05
step: 390, loss: 6.588001997442916e-05
step: 400, loss: 0.0004179595271125436
step: 410, loss: 6.8118424678687e-05
step: 420, loss: 4.4363780034473166e-05
step: 430, loss: 3.409734199522063e-05
step: 440, loss: 8.076781523413956e-05
step: 450, loss: 0.0052860495634377
step: 460, loss: 0.0033103099558502436
step: 470, loss: 7.977543282322586e-05
step: 480, loss: 0.002016331534832716
step: 490, loss: 0.0004226517630741
step: 500, loss: 0.0008727000094950199
step: 510, loss: 7.306714542210102e-05
step: 520, loss: 0.00016918468463700265
step: 530, loss: 4.436639937921427e-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=0.9530887134231305, f1=0.9500462534690102,
best_f1=0.948356807511737
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:34, 166.19it/s]
load_f1 = 0.9575361642557162
real_f1 = 0.9589169000933706
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.71it/s]

```

4 DITTO

4.1 DITTO STRUCTURED

4.1.1 Beer - Running the matcher

```

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

```

```

--task Structured/Beer \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.4620782136917114
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.2947368421052632, f1=0.27184466019417475,
best_f1=0.27184466019417475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4355957508087158
/usr/local/lib/python3.7/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=0.2947368421052632, f1=0.2692307692307693,
best_f1=0.27184466019417475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3897997736930847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.48, f1=0.35294117647058826, best_f1=0.35294117647058826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28581491112709045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5217391304347825, f1=0.46153846153846156,
best_f1=0.46153846153846156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24815836548805237
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.6666666666666665,
best_f1=0.6666666666666665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15010076761245728
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.8275862068965518, f1=0.7222222222222223,
best_f1=0.6666666666666665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27455177903175354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.896551724137931, f1=0.7333333333333334,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16902495920658112
/usr/local/lib/python3.7/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=0.8235294117647058, f1=0.717948717948718,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1791723519563675
/usr/local/lib/python3.7/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=0.8888888888888889, f1=0.7200000000000001,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16954712569713593
/usr/local/lib/python3.7/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=0.8888888888888889, f1=0.7692307692307692,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0475720539689064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8666666666666666, f1=0.9032258064516129,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016384422779083252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8666666666666666, f1=0.8484848484848484,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012074840255081654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8666666666666666, f1=0.8750000000000001,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006889475509524345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8666666666666666, f1=0.8750000000000001,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012365891598165035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8666666666666666, f1=0.8750000000000001,
best_f1=0.7333333333333334
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 85253.89it/s]
/usr/local/lib/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 = 0.7857142857142857
real_f1 = 0.8387096774193549
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 142.78it/s]

```

4.1.2 DBLP-ACM - Running the matcher

```

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

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

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

```

```
--output_path output/output_small.jsonl \  
--lm roberta \  
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

```
Downloading: 100% 481/481 [00:00<00:00, 415kB/s]  
Downloading: 100% 899k/899k [00:00<00:00, 6.04MB/s]  
Downloading: 100% 456k/456k [00:00<00:00, 2.96MB/s]  
Downloading: 100% 501M/501M [00:07<00:00, 63.0MB/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: 0.5571184754371643

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

step: 10, loss: 0.4547971785068512

step: 20, loss: 0.5089476108551025

step: 30, loss: 0.326252281665802

step: 40, loss: 0.3525920808315277

step: 50, loss: 0.43020012974739075

step: 60, loss: 0.13975277543067932

step: 70, loss: 0.09474581480026245

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

step: 80, loss: 0.210772305727005

step: 90, loss: 0.14980663359165192

step: 100, loss: 0.0712539404630661

step: 110, loss: 0.026107678189873695

step: 120, loss: 0.06513062119483948

step: 130, loss: 0.014936202205717564

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

step: 140, loss: 0.12716346979141235

step: 150, loss: 0.09087912738323212

step: 160, loss: 0.01703525334596634

step: 170, loss: 0.035207897424697876

step: 180, loss: 0.04009120166301727

step: 190, loss: 0.026902396231889725

step: 200, loss: 0.17708715796470642

step: 210, loss: 0.07163768261671066

step: 220, loss: 0.029884839430451393

step: 230, loss: 0.1227019727230072

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

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9596510359869138, f1=0.9551912568306011,
best_f1=0.9551912568306011

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

step: 0, loss: 0.0056521412916481495

step: 10, loss: 0.06425651162862778

step: 20, loss: 0.01375807449221611

step: 30, loss: 0.02596341073513031

step: 40, loss: 0.00403363024815917

step: 50, loss: 0.0017544366419315338

step: 60, loss: 0.005841127596795559

step: 70, loss: 0.026557806879281998

step: 80, loss: 0.003974668215960264

step: 90, loss: 0.0046739825047552586

```

step: 100, loss: 0.004343666601926088
step: 110, loss: 0.0038396401796489954
step: 120, loss: 0.00966495368629694
step: 130, loss: 0.010035661049187183
step: 140, loss: 0.026336681097745895
step: 150, loss: 0.16441752016544342
step: 160, loss: 0.010978887788951397
step: 170, loss: 0.027699556201696396
step: 180, loss: 0.007507207337766886
step: 190, loss: 0.09999266266822815
step: 200, loss: 0.007357250899076462
step: 210, loss: 0.0065422882325947285
step: 220, loss: 0.0009187002433463931
step: 230, loss: 0.0008606495684944093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9809203142536477, f1=0.9752808988764046,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006977270822972059
step: 10, loss: 0.008546289056539536
step: 20, loss: 0.03132907673716545
step: 30, loss: 0.0009501964086666703
step: 40, loss: 0.006407429929822683
step: 50, loss: 0.07900890707969666
step: 60, loss: 0.002140816766768694
step: 70, loss: 0.028745412826538086
step: 80, loss: 0.001043857540935278
step: 90, loss: 0.019408343359827995
step: 100, loss: 0.009219052270054817
step: 110, loss: 0.0015479808207601309
step: 120, loss: 0.0008105872548185289
step: 130, loss: 0.0014952468918636441
step: 140, loss: 0.00150591554120183
step: 150, loss: 0.01892905868589878
step: 160, loss: 0.0020774570293724537
step: 170, loss: 0.0011529568582773209
step: 180, loss: 0.006165739148855209
step: 190, loss: 0.040015485137701035
step: 200, loss: 0.015825865790247917
step: 210, loss: 0.0015911075752228498

```

```

step: 220, loss: 0.004231204744428396
step: 230, loss: 0.008059952408075333
/usr/local/lib/python3.7/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=0.9741282339707535, f1=0.9751693002257337,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013377897441387177
step: 10, loss: 0.03867235779762268
step: 20, loss: 0.009113507345318794
step: 30, loss: 0.0026857987977564335
step: 40, loss: 0.07505045086145401
step: 50, loss: 0.02889716625213623
step: 60, loss: 0.0021421120036393404
step: 70, loss: 0.0057172649540007114
step: 80, loss: 0.0006819369737058878
step: 90, loss: 0.05572260171175003
step: 100, loss: 0.005016495008021593
step: 110, loss: 0.0004709097556769848
step: 120, loss: 0.0011975881643593311
step: 130, loss: 0.018140651285648346
step: 140, loss: 0.00038166515878401697
step: 150, loss: 0.0004175025096628815
step: 160, loss: 0.000444899225840345
step: 170, loss: 0.05735538899898529
step: 180, loss: 0.13349150121212006
step: 190, loss: 0.0033524902537465096
step: 200, loss: 0.038366757333278656
step: 210, loss: 0.0018654598388820887
step: 220, loss: 0.0008091090712696314
step: 230, loss: 0.00241085235029459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9842696629213483, f1=0.9853107344632768,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001152229611761868
step: 10, loss: 0.02755187824368477

```



```

step: 20, loss: 0.021301133558154106
step: 30, loss: 0.0007139646331779659
step: 40, loss: 0.0009433482773602009
step: 50, loss: 0.0024546703789383173
step: 60, loss: 0.0023304447531700134
step: 70, loss: 0.0006627184920944273
step: 80, loss: 0.00918598286807537
step: 90, loss: 0.06554576009511948
step: 100, loss: 0.0006398437544703484
step: 110, loss: 0.00048038209206424654
step: 120, loss: 0.003939863760024309
step: 130, loss: 0.00041748667717911303
step: 140, loss: 0.0014255708083510399
step: 150, loss: 0.02731780894100666
step: 160, loss: 0.001101775444112718
step: 170, loss: 0.003941123373806477
step: 180, loss: 0.0008150412468239665
step: 190, loss: 0.11710339784622192
step: 200, loss: 0.02210455946624279
step: 210, loss: 0.025139786303043365
step: 220, loss: 0.0015924365725368261
step: 230, loss: 0.003547766013070941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854423292273236, f1=0.9841986455981941,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007830924587324262
step: 10, loss: 0.0017798448679968715
step: 20, loss: 0.00141806248575449
step: 30, loss: 0.005091627594083548
step: 40, loss: 0.000475464592454955
step: 50, loss: 0.0005190116353332996
step: 60, loss: 0.04627177119255066
step: 70, loss: 0.0028734535444527864
step: 80, loss: 0.020281950011849403
step: 90, loss: 0.0184311643242836
step: 100, loss: 0.0006286168936640024
step: 110, loss: 0.039056241512298584
step: 120, loss: 0.00020884719560854137
step: 130, loss: 0.010018006898462772

```

```

step: 140, loss: 0.00024799490347504616
step: 150, loss: 0.007183757144957781
step: 160, loss: 0.017647387459874153
step: 170, loss: 0.0003112497797701508
step: 180, loss: 0.0008674376294948161
step: 190, loss: 0.002748744795098901
step: 200, loss: 0.056238219141960144
step: 210, loss: 0.02440202794969082
step: 220, loss: 0.006509333848953247
step: 230, loss: 0.0014263774501159787
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9852774631936579,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00281129009090364
step: 10, loss: 0.00244984426535666
step: 20, loss: 0.0038927439600229263
step: 30, loss: 0.0018283601384609938
step: 40, loss: 0.003142197849228978
step: 50, loss: 0.0006086511421017349
step: 60, loss: 0.00021295329497661442
step: 70, loss: 0.00017652598035056144
step: 80, loss: 0.0008298372849822044
step: 90, loss: 0.07933753728866577
step: 100, loss: 0.0009421510621905327
step: 110, loss: 0.0005270606488920748
step: 120, loss: 0.002016794867813587
step: 130, loss: 0.0031714087817817926
step: 140, loss: 0.0006583920912817121
step: 150, loss: 0.013137124478816986
step: 160, loss: 0.0008367023547179997
step: 170, loss: 0.020539700984954834
step: 180, loss: 0.0007127450662665069
step: 190, loss: 0.0003582545614335686
step: 200, loss: 0.01257602870464325
step: 210, loss: 0.08922228217124939
step: 220, loss: 0.0006515348213724792
step: 230, loss: 0.007610222790390253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887640449438202, f1=0.9864559819413092,
best_f1=0.9864559819413092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005448139738291502
step: 10, loss: 0.003050890751183033
step: 20, loss: 0.000932756345719099
step: 30, loss: 0.0005268182139843702
step: 40, loss: 0.028185244649648666
step: 50, loss: 0.0008469106396660209
step: 60, loss: 0.0007340650190599263
step: 70, loss: 0.0002119774726452306
step: 80, loss: 0.03455161303281784
step: 90, loss: 0.0005734125734306872
step: 100, loss: 0.0006106326472945511
step: 110, loss: 0.0002901660627685487
step: 120, loss: 0.0004644187865778804
step: 130, loss: 0.0036196173168718815
step: 140, loss: 0.0007525988039560616
step: 150, loss: 0.04890598729252815
step: 160, loss: 0.0051281931810081005
step: 170, loss: 0.06325550377368927
step: 180, loss: 0.0004644868313334882
step: 190, loss: 0.006105493288487196
step: 200, loss: 0.0037512679118663073
step: 210, loss: 0.0034380434080958366
step: 220, loss: 0.0004930155701003969
step: 230, loss: 0.0003075927961617708
/usr/local/lib/python3.7/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=0.987736900780379, f1=0.9810055865921787,
best_f1=0.9864559819413092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005357021582312882
step: 10, loss: 0.0005307900719344616
step: 20, loss: 0.0011159003479406238
step: 30, loss: 0.0006866385228931904
step: 40, loss: 0.002860149135813117
step: 50, loss: 0.0006506055360659957
step: 60, loss: 0.00016089485143311322
step: 70, loss: 0.025853270664811134
step: 80, loss: 0.00040461920434609056
step: 90, loss: 0.04620451480150223

```

```

step: 100, loss: 0.000157383838086389
step: 110, loss: 0.0004456256574485451
step: 120, loss: 0.07638388872146606
step: 130, loss: 0.0005516470409929752
step: 140, loss: 0.0037901420146226883
step: 150, loss: 0.0013709368649870157
step: 160, loss: 0.0005019669188186526
step: 170, loss: 0.0009973908308893442
step: 180, loss: 0.0001909861748572439
step: 190, loss: 9.821398271014914e-05
step: 200, loss: 0.027429156005382538
step: 210, loss: 0.032339222729206085
step: 220, loss: 0.00023952455376274884
step: 230, loss: 0.0031267309095710516
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9864559819413092, f1=0.9819413092550789,
best_f1=0.9864559819413092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029541627736762166
step: 10, loss: 0.0007145173149183393
step: 20, loss: 0.0006885498878546059
step: 30, loss: 0.0001537794159958139
step: 40, loss: 0.004032955504953861
step: 50, loss: 4.9405462050344795e-05
step: 60, loss: 5.742630310123786e-05
step: 70, loss: 0.053021389991045
step: 80, loss: 4.3584295781329274e-05
step: 90, loss: 6.388096517184749e-05
step: 100, loss: 5.7911289331968874e-05
step: 110, loss: 0.00534376036375761
step: 120, loss: 7.403617928503081e-05
step: 130, loss: 0.00014194240793585777
step: 140, loss: 7.078640919644386e-05
step: 150, loss: 0.009145648218691349
step: 160, loss: 3.477298014331609e-05
step: 170, loss: 6.167539686430246e-05
step: 180, loss: 0.00011614546383498237
step: 190, loss: 0.00010163422848563641
step: 200, loss: 8.032277401071042e-05
step: 210, loss: 0.00021341275714803487
step: 220, loss: 7.228289177874103e-05
step: 230, loss: 0.0005872318870387971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9887640449438202, f1=0.9831271091113611,
best_f1=0.9864559819413092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036772829480469227
step: 10, loss: 8.074907964328304e-05
step: 20, loss: 5.8474917750572786e-05
step: 30, loss: 6.683314859401435e-05
step: 40, loss: 2.7856398446601816e-05
step: 50, loss: 0.003158334642648697
step: 60, loss: 0.009408269077539444
step: 70, loss: 0.0057839443907141685
step: 80, loss: 0.0003454993129707873
step: 90, loss: 0.06906718015670776
step: 100, loss: 0.0007054173038341105
step: 110, loss: 0.0006716093048453331
step: 120, loss: 0.00015349585737567395
step: 130, loss: 0.00019219121895730495
step: 140, loss: 0.013123851269483566
step: 150, loss: 0.00021430719061754644
step: 160, loss: 0.06969203054904938
step: 170, loss: 0.001298008719459176
step: 180, loss: 0.00021639444457832724
step: 190, loss: 0.0001253537047887221
step: 200, loss: 0.004833343904465437
step: 210, loss: 7.379087765002623e-05
step: 220, loss: 0.0011218705913051963
step: 230, loss: 0.04832775145769119
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9808773903262092,
best_f1=0.9864559819413092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011874970368808135
step: 10, loss: 0.00010407741501694545
step: 20, loss: 0.008841698057949543
step: 30, loss: 0.020194925367832184
step: 40, loss: 0.0002884138375520706
step: 50, loss: 0.0013841385953128338
step: 60, loss: 0.013640346936881542
step: 70, loss: 0.00017810142890084535
step: 80, loss: 4.5995668187970296e-05
step: 90, loss: 0.00235801562666893

```

```

step: 100, loss: 5.3374082199297845e-05
step: 110, loss: 4.1403323848498985e-05
step: 120, loss: 0.0002692873531486839
step: 130, loss: 5.291322304401547e-05
step: 140, loss: 8.005223207874224e-05
step: 150, loss: 0.0007001247722655535
step: 160, loss: 0.0001250748900929466
step: 170, loss: 0.0028962434735149145
step: 180, loss: 5.3398849559016526e-05
step: 190, loss: 0.006834044586867094
step: 200, loss: 4.517148408922367e-05
step: 210, loss: 0.00796093512326479
step: 220, loss: 0.005784039851278067
step: 230, loss: 7.33316846890375e-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=0.9910313901345291, f1=0.9810479375696767,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011173069477081299
step: 10, loss: 0.0009749981691129506
step: 20, loss: 0.00011271137918811291
step: 30, loss: 0.0030809538438916206
step: 40, loss: 6.244029646040872e-05
step: 50, loss: 0.0005644949851557612
step: 60, loss: 0.0005551881622523069
step: 70, loss: 0.03337322548031807
step: 80, loss: 0.00020813824085053056
step: 90, loss: 0.0001163445194833912
step: 100, loss: 0.0001668213080847636
step: 110, loss: 0.000717436196282506
step: 120, loss: 0.00014867396384943277
step: 130, loss: 0.00011377024202374741
step: 140, loss: 0.00017528163152746856
step: 150, loss: 8.733115828363225e-05
step: 160, loss: 0.009903962723910809
step: 170, loss: 6.905738700879738e-05
step: 180, loss: 0.03992258012294769
step: 190, loss: 0.00014366628602147102
step: 200, loss: 0.00011291939881630242
step: 210, loss: 0.00024747051065787673

```

```

step: 220, loss: 0.00020220738952048123
step: 230, loss: 7.538011413998902e-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=0.990990990990991, f1=0.9819413092550789,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.576883293339051e-05
step: 10, loss: 7.296170952031389e-05
step: 20, loss: 8.534798689652234e-05
step: 30, loss: 0.0031314275693148375
step: 40, loss: 0.0005225134082138538
step: 50, loss: 3.583440411603078e-05
step: 60, loss: 0.00011536631791386753
step: 70, loss: 9.564743959344923e-05
step: 80, loss: 5.559671626542695e-05
step: 90, loss: 4.845027069677599e-05
step: 100, loss: 0.0005150371580384672
step: 110, loss: 5.647432044497691e-05
step: 120, loss: 2.138990930689033e-05
step: 130, loss: 7.976108463481069e-05
step: 140, loss: 3.802963692578487e-05
step: 150, loss: 7.879883196437731e-05
step: 160, loss: 0.0006694091134704649
step: 170, loss: 0.0001571247266838327
step: 180, loss: 3.9478061808040366e-05
step: 190, loss: 0.001240517827682197
step: 200, loss: 5.166221671970561e-05
step: 210, loss: 6.222595402505249e-05
step: 220, loss: 7.668867328902707e-05
step: 230, loss: 0.0004705112660303712
/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9830890642615557,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0318833589553833
step: 10, loss: 4.216047454974614e-05
step: 20, loss: 0.000341554987244308
step: 30, loss: 7.411393744405359e-05
step: 40, loss: 0.0005230465903878212
step: 50, loss: 3.810758062172681e-05

```

```

step: 60, loss: 0.02587631531059742
step: 70, loss: 0.00010884339280892164
step: 80, loss: 3.990623008576222e-05
step: 90, loss: 8.324460941366851e-05
step: 100, loss: 0.00012394173245411366
step: 110, loss: 0.00012033598613925278
step: 120, loss: 0.0219047162681818
step: 130, loss: 0.00014558089605998248
step: 140, loss: 0.004634020384401083
step: 150, loss: 8.878044172888622e-05
step: 160, loss: 0.017741771414875984
step: 170, loss: 2.1863086658413522e-05
step: 180, loss: 0.00023138443066272885
step: 190, loss: 6.982870399951935e-05
step: 200, loss: 0.00030455130035988986
step: 210, loss: 0.06005102023482323
step: 220, loss: 4.727043415186927e-05
step: 230, loss: 0.0001414291764376685
/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9842342342342343,
best_f1=0.9810479375696767
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:12, 205.71it/s]
load_f1 = 0.9887892376681614
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 [00:23, 185.95it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.6003347039222717
```

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

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

```
step: 10, loss: 0.44009873270988464
```

```
step: 20, loss: 0.23196345567703247
```

```
step: 30, loss: 0.384991854429245
```

```
step: 40, loss: 0.3761900067329407
```

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

```
step: 50, loss: 0.49571672081947327
```

```
step: 60, loss: 0.33929967880249023
```

```
step: 70, loss: 0.13385532796382904
```

```
step: 80, loss: 0.19803635776042938
```

```
step: 90, loss: 0.1756998896598816
```

```
step: 100, loss: 0.1591089516878128
```

```
step: 110, loss: 0.1710689216852188
```

```
step: 120, loss: 0.12851016223430634
```

```
step: 130, loss: 0.09186941385269165
```

```
step: 140, loss: 0.2103874236345291
```

```

step: 150, loss: 0.11782488226890564
step: 160, loss: 0.1561291366815567
step: 170, loss: 0.017030376940965652
step: 180, loss: 0.06858668476343155
step: 190, loss: 0.0901590958237648
step: 200, loss: 0.01589100994169712
step: 210, loss: 0.02339087426662445
step: 220, loss: 0.07939070463180542
step: 230, loss: 0.1869005709886551
step: 240, loss: 0.06868265569210052
step: 250, loss: 0.017085928469896317
step: 260, loss: 0.317659854888916
step: 270, loss: 0.22827698290348053
step: 280, loss: 0.08762544393539429
step: 290, loss: 0.03568250313401222
step: 300, loss: 0.08086016029119492
step: 310, loss: 0.13949653506278992
step: 320, loss: 0.07840052247047424
step: 330, loss: 0.047867294400930405
step: 340, loss: 0.4678080976009369
step: 350, loss: 0.17529097199440002
step: 360, loss: 0.012781822122633457
step: 370, loss: 0.01924186199903488
step: 380, loss: 0.060081347823143005
step: 390, loss: 0.029125714674592018
step: 400, loss: 0.054593656212091446
step: 410, loss: 0.26033249497413635
step: 420, loss: 0.03550586476922035
step: 430, loss: 0.05540226399898529
step: 440, loss: 0.030287278816103935
step: 450, loss: 0.02072971500456333
step: 460, loss: 0.01386328600347042
step: 470, loss: 0.017258398234844208
step: 480, loss: 0.1495843529701233
step: 490, loss: 0.14651800692081451
step: 500, loss: 0.07022316753864288
step: 510, loss: 0.13961927592754364
step: 520, loss: 0.32840636372566223
step: 530, loss: 0.017873723059892654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9345622119815669, f1=0.9332105020727776,
best_f1=0.9332105020727776

```

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

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

```
step: 0, loss: 0.09339568018913269  
step: 10, loss: 0.08985365927219391  
step: 20, loss: 0.04790868982672691  
step: 30, loss: 0.050902679562568665  
step: 40, loss: 0.05334793031215668  
step: 50, loss: 0.022748520597815514  
step: 60, loss: 0.08535590767860413  
step: 70, loss: 0.07728558033704758  
step: 80, loss: 0.018342992290854454  
step: 90, loss: 0.006353222299367189  
step: 100, loss: 0.10473329573869705  
step: 110, loss: 0.011073512025177479  
step: 120, loss: 0.14829443395137787  
step: 130, loss: 0.013387330807745457  
step: 140, loss: 0.11902576684951782  
step: 150, loss: 0.006839197129011154  
step: 160, loss: 0.04018162190914154  
step: 170, loss: 0.11308873444795609  
step: 180, loss: 0.031195281073451042  
step: 190, loss: 0.03316260501742363  
step: 200, loss: 0.17854733765125275  
step: 210, loss: 0.024588599801063538  
step: 220, loss: 0.0033879803959280252  
step: 230, loss: 0.012725789099931717  
step: 240, loss: 0.023082418367266655  
step: 250, loss: 0.08190261572599411  
step: 260, loss: 0.05221843346953392  
step: 270, loss: 0.025310475379228592  
step: 280, loss: 0.14571379125118256  
step: 290, loss: 0.07652604579925537  
step: 300, loss: 0.03967166319489479  
step: 310, loss: 0.1805645078420639  
step: 320, loss: 0.05879371240735054  
step: 330, loss: 0.11620617657899857  
step: 340, loss: 0.03335035592317581  
step: 350, loss: 0.004536710679531097  
step: 360, loss: 0.052114903926849365  
step: 370, loss: 0.028320306912064552  
step: 380, loss: 0.14524634182453156  
step: 390, loss: 0.006464356556534767  
step: 400, loss: 0.02606823481619358  
step: 410, loss: 0.025448691099882126  
step: 420, loss: 0.059724658727645874  
step: 430, loss: 0.1422087699174881  
step: 440, loss: 0.03391382470726967
```

```

step: 450, loss: 0.15872539579868317
step: 460, loss: 0.05175276845693588
step: 470, loss: 0.051638782024383545
step: 480, loss: 0.02114223875105381
step: 490, loss: 0.04144136607646942
step: 500, loss: 0.009319709613919258
step: 510, loss: 0.014262969605624676
step: 520, loss: 0.3283609449863434
step: 530, loss: 0.016314508393406868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9477093937991671, f1=0.9443929564411491,
best_f1=0.9443929564411491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1502699851989746
step: 10, loss: 0.017738858237862587
step: 20, loss: 0.011924118734896183
step: 30, loss: 0.1790846586227417
step: 40, loss: 0.019994819536805153
step: 50, loss: 0.00654454855248332
step: 60, loss: 0.06804290413856506
step: 70, loss: 0.023174427449703217
step: 80, loss: 0.04333172366023064
step: 90, loss: 0.01902671530842781
step: 100, loss: 0.00784558430314064
step: 110, loss: 0.026581896468997
step: 120, loss: 0.1662011295557022
step: 130, loss: 0.037127505987882614
step: 140, loss: 0.026834925636649132
step: 150, loss: 0.04891154542565346
step: 160, loss: 0.012770306318998337
step: 170, loss: 0.009581348858773708
step: 180, loss: 0.02234017476439476
step: 190, loss: 0.019641349092125893
step: 200, loss: 0.1506042182445526
step: 210, loss: 0.042372968047857285
step: 220, loss: 0.19054532051086426
step: 230, loss: 0.03574439138174057
step: 240, loss: 0.02593330293893814
step: 250, loss: 0.08641500025987625
step: 260, loss: 0.10760756582021713

```

```

step: 270, loss: 0.008356552571058273
step: 280, loss: 0.00280699972063303
step: 290, loss: 0.002501150593161583
step: 300, loss: 0.07237143814563751
step: 310, loss: 0.05719583109021187
step: 320, loss: 0.011796497739851475
step: 330, loss: 0.014285448007285595
step: 340, loss: 0.01252404972910881
step: 350, loss: 0.12123265117406845
step: 360, loss: 0.02628142759203911
step: 370, loss: 0.02554069645702839
step: 380, loss: 0.048527587205171585
step: 390, loss: 0.009687505662441254
step: 400, loss: 0.08854657411575317
step: 410, loss: 0.02260543406009674
step: 420, loss: 0.06470783054828644
step: 430, loss: 0.03288492187857628
step: 440, loss: 0.22123867273330688
step: 450, loss: 0.0855429619550705
step: 460, loss: 0.138068288564682
step: 470, loss: 0.02882196754217148
step: 480, loss: 0.09050162136554718
step: 490, loss: 0.020007794722914696
step: 500, loss: 0.009399175643920898
step: 510, loss: 0.022662591189146042
step: 520, loss: 0.013032921589910984
step: 530, loss: 0.003272583708167076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9512867647058824, f1=0.9472718936267768,
best_f1=0.9472718936267768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024933168664574623
step: 10, loss: 0.003599008545279503
step: 20, loss: 0.1870446652173996
step: 30, loss: 0.05323387309908867
step: 40, loss: 0.014209180139005184
step: 50, loss: 0.06401511281728745
step: 60, loss: 0.012809846550226212
step: 70, loss: 0.008193755522370338
step: 80, loss: 0.0028319652192294598

```

step: 90, loss: 0.026979610323905945
step: 100, loss: 0.0014948549214750528
step: 110, loss: 0.007815110497176647
step: 120, loss: 0.025906410068273544
step: 130, loss: 0.007865617983043194
step: 140, loss: 0.01345906313508749
step: 150, loss: 0.006615910213440657
step: 160, loss: 0.02133788913488388
step: 170, loss: 0.017859235405921936
step: 180, loss: 0.1260865181684494
step: 190, loss: 0.0490606315433979
step: 200, loss: 0.04750797525048256
step: 210, loss: 0.002893191995099187
step: 220, loss: 0.005175345577299595
step: 230, loss: 0.005152999889105558
step: 240, loss: 0.0011057500960305333
step: 250, loss: 0.1455824375152588
step: 260, loss: 0.0017446039710193872
step: 270, loss: 0.04534219205379486
step: 280, loss: 0.0064215390011668205
step: 290, loss: 0.02484455332159996
step: 300, loss: 0.018901994451880455
step: 310, loss: 0.0006705619161948562
step: 320, loss: 0.04115325212478638
step: 330, loss: 0.06565925478935242
step: 340, loss: 0.053488388657569885
step: 350, loss: 0.08232199400663376
step: 360, loss: 0.034572288393974304
step: 370, loss: 0.010575509630143642
step: 380, loss: 0.03074842318892479
step: 390, loss: 0.0016857597511261702
step: 400, loss: 0.012884721159934998
step: 410, loss: 0.01683911681175232
step: 420, loss: 0.018877187743782997
step: 430, loss: 0.009327530860900879
step: 440, loss: 0.022603826597332954
step: 450, loss: 0.012707713060081005
step: 460, loss: 0.003829309018328786
step: 470, loss: 0.0021246757823973894
step: 480, loss: 0.011172437109053135
step: 490, loss: 0.005537834018468857
step: 500, loss: 0.0988599956035614
step: 510, loss: 0.08144834637641907
step: 520, loss: 0.04277750477194786
step: 530, loss: 0.035409145057201385

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 save or load the state of the optimizer when saving or
loading the scheduler.
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9550561797752809, f1=0.9489939167056621,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009825410088524222
step: 10, loss: 0.008835206739604473
step: 20, loss: 0.004140134435147047
step: 30, loss: 0.006376477424055338
step: 40, loss: 0.0016157242935150862
step: 50, loss: 0.11415305733680725
step: 60, loss: 0.05678332597017288
step: 70, loss: 0.02341138757765293
step: 80, loss: 0.049766600131988525
step: 90, loss: 0.06104409694671631
step: 100, loss: 0.10378540307283401
step: 110, loss: 0.028786929324269295
step: 120, loss: 0.17806454002857208
step: 130, loss: 0.007034025155007839
step: 140, loss: 0.017431480810046196
step: 150, loss: 0.010458534583449364
step: 160, loss: 0.029914673417806625
step: 170, loss: 0.07279708981513977
step: 180, loss: 0.008226446807384491
step: 190, loss: 0.007597512099891901
step: 200, loss: 0.0015167332021519542
step: 210, loss: 0.0006514484994113445
step: 220, loss: 0.022814916446805
step: 230, loss: 0.01435638964176178
step: 240, loss: 0.008904571644961834
step: 250, loss: 0.09026303142309189
step: 260, loss: 0.0030059083364903927
step: 270, loss: 0.18246133625507355
step: 280, loss: 0.011988168582320213
step: 290, loss: 0.017170852050185204
step: 300, loss: 0.2600690722465515
step: 310, loss: 0.020958106964826584
step: 320, loss: 0.01832025870680809
step: 330, loss: 0.0017155511304736137
step: 340, loss: 0.01306107360869646
step: 350, loss: 0.0006590025732293725
step: 360, loss: 0.00025066095986403525
step: 370, loss: 0.0008974300581030548

```



```
step: 380, loss: 0.004505994729697704
step: 390, loss: 0.0009615116869099438
step: 400, loss: 0.003477443242445588
step: 410, loss: 0.06651009619235992
step: 420, loss: 0.20022737979888916
step: 430, loss: 0.04645255208015442
step: 440, loss: 0.004102885257452726
step: 450, loss: 0.0032368518877774477
step: 460, loss: 0.01298394612967968
step: 470, loss: 0.020112480968236923
step: 480, loss: 0.24357587099075317
step: 490, loss: 0.04045308381319046
step: 500, loss: 0.11168044805526733
step: 510, loss: 0.01930098608136177
step: 520, loss: 0.09329552948474884
step: 530, loss: 0.058120571076869965
/usr/local/lib/python3.7/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=0.954119850187266, f1=0.9438202247191012,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.194025456905365
step: 10, loss: 0.0034091342240571976
step: 20, loss: 0.038797520101070404
step: 30, loss: 0.0036455329973250628
step: 40, loss: 0.05067499727010727
step: 50, loss: 0.03654770180583
step: 60, loss: 0.012171749025583267
step: 70, loss: 0.00343883503228426
step: 80, loss: 0.0012412999058142304
step: 90, loss: 0.13372543454170227
step: 100, loss: 0.0027709933929145336
step: 110, loss: 0.004271623212844133
step: 120, loss: 0.03710530325770378
step: 130, loss: 0.004188838880509138
step: 140, loss: 0.0035843243822455406
step: 150, loss: 0.0012972477125003934
step: 160, loss: 0.040992122143507004
step: 170, loss: 0.003241756232455373
step: 180, loss: 0.002939121797680855
step: 190, loss: 0.216959148645401
step: 200, loss: 0.01330883800983429
step: 210, loss: 0.00319383735768497
step: 220, loss: 0.048217952251434326
step: 230, loss: 0.0017975657247006893
```

```

step: 240, loss: 0.0008656203863210976
step: 250, loss: 0.0380004420876503
step: 260, loss: 0.004827044904232025
step: 270, loss: 0.0011224234476685524
step: 280, loss: 0.0023138057440519333
step: 290, loss: 0.0003689885197672993
step: 300, loss: 0.00393203878775239
step: 310, loss: 0.1466132402420044
step: 320, loss: 0.0004244843148626387
step: 330, loss: 0.0007498508202843368
step: 340, loss: 0.003965418785810471
step: 350, loss: 0.01350912731140852
step: 360, loss: 0.013709161430597305
step: 370, loss: 0.0281665176153183
step: 380, loss: 0.001811079098843038
step: 390, loss: 0.001737948041409254
step: 400, loss: 0.19862952828407288
step: 410, loss: 0.01821267046034336
step: 420, loss: 0.002744824392721057
step: 430, loss: 0.000750939769204706
step: 440, loss: 0.006440326105803251
step: 450, loss: 0.1871042251586914
step: 460, loss: 0.0027055132668465376
step: 470, loss: 0.001858277479186654
step: 480, loss: 0.021288124844431877
step: 490, loss: 0.02960534766316414
step: 500, loss: 0.0004895589081570506
step: 510, loss: 0.13847461342811584
step: 520, loss: 0.002240309491753578
step: 530, loss: 0.0002810674486681819
/usr/local/lib/python3.7/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=0.9547363509099394, f1=0.9445738239403819,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03808574378490448
step: 10, loss: 0.00080119539052248
step: 20, loss: 0.007596948184072971
step: 30, loss: 0.01093664113432169
step: 40, loss: 0.002070200862362981
step: 50, loss: 0.00367534882389009
step: 60, loss: 0.03271147236227989
step: 70, loss: 0.028530068695545197
step: 80, loss: 0.0007495615282095969
step: 90, loss: 0.0003495600540190935

```

```
step: 100, loss: 0.003196598030626774
step: 110, loss: 0.00021321122767403722
step: 120, loss: 0.0005431136232800782
step: 130, loss: 0.0016594022745266557
step: 140, loss: 0.004144384991377592
step: 150, loss: 0.003632543608546257
step: 160, loss: 0.0006853851373307407
step: 170, loss: 0.007780815474689007
step: 180, loss: 0.08222349733114243
step: 190, loss: 0.12271318584680557
step: 200, loss: 0.00617055082693696
step: 210, loss: 0.07207228988409042
step: 220, loss: 0.05654924735426903
step: 230, loss: 0.00018314090266358107
step: 240, loss: 0.018548427149653435
step: 250, loss: 0.008769973181188107
step: 260, loss: 0.11453146487474442
step: 270, loss: 0.0019984813407063484
step: 280, loss: 0.021220451220870018
step: 290, loss: 0.0006445567705668509
step: 300, loss: 0.0005015717470087111
step: 310, loss: 0.005575982388108969
step: 320, loss: 0.09913478046655655
step: 330, loss: 0.0016032805433496833
step: 340, loss: 0.004118679091334343
step: 350, loss: 0.008732292801141739
step: 360, loss: 0.030300159007310867
step: 370, loss: 0.05266512930393219
step: 380, loss: 0.08975031971931458
step: 390, loss: 0.050630874931812286
step: 400, loss: 0.06287669390439987
step: 410, loss: 0.021077826619148254
step: 420, loss: 0.005557623226195574
step: 430, loss: 0.00039597766590304673
step: 440, loss: 0.0008218815783038735
step: 450, loss: 0.010897113010287285
step: 460, loss: 0.003852268448099494
step: 470, loss: 0.12088098376989365
step: 480, loss: 0.009906991384923458
step: 490, loss: 0.0011144025484099984
step: 500, loss: 0.0031504963990300894
step: 510, loss: 0.000253954523941502
step: 520, loss: 0.0010976827470585704
step: 530, loss: 0.0010283497394993901
/usr/local/lib/python3.7/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=0.9500699953336444, f1=0.947565543071161,
```

```
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021611833944916725
step: 10, loss: 0.006324792746454477
step: 20, loss: 0.0012615332379937172
step: 30, loss: 0.0017947499873116612
step: 40, loss: 0.00027490584761835635
step: 50, loss: 0.00015811367484275252
step: 60, loss: 0.0003361851559020579
step: 70, loss: 0.0022623410914093256
step: 80, loss: 0.009011251851916313
step: 90, loss: 0.00042912494973279536
step: 100, loss: 0.012136452831327915
step: 110, loss: 0.0005013309419155121
step: 120, loss: 0.0026858914643526077
step: 130, loss: 0.007491614669561386
step: 140, loss: 7.574685150757432e-05
step: 150, loss: 0.0003228116547688842
step: 160, loss: 6.144013605080545e-05
step: 170, loss: 0.02194204553961754
step: 180, loss: 0.0002581071457825601
step: 190, loss: 0.0007780360756441951
step: 200, loss: 0.003943123389035463
step: 210, loss: 0.05922279134392738
step: 220, loss: 0.0018362945411354303
step: 230, loss: 0.0487089566886425
step: 240, loss: 0.009493347257375717
step: 250, loss: 0.009239018894731998
step: 260, loss: 0.0020818666089326143
step: 270, loss: 0.008590423502027988
step: 280, loss: 0.0006049475632607937
step: 290, loss: 0.0010489514097571373
step: 300, loss: 2.8501488486654125e-05
step: 310, loss: 0.00143647741060704
step: 320, loss: 0.001484779641032219
step: 330, loss: 0.0002389859000686556
step: 340, loss: 0.17660710215568542
step: 350, loss: 0.0008778587216511369
step: 360, loss: 0.031835976988077164
step: 370, loss: 0.019013304263353348
step: 380, loss: 0.002757543232291937
step: 390, loss: 0.006647492293268442
step: 400, loss: 0.0021095722913742065
step: 410, loss: 0.0002932960633188486
step: 420, loss: 7.23047778592445e-05
step: 430, loss: 0.1055125892162323
```

```

step: 440, loss: 0.03378206491470337
step: 450, loss: 0.0016220908146351576
step: 460, loss: 0.006502749398350716
step: 470, loss: 0.08273053169250488
step: 480, loss: 0.018124613910913467
step: 490, loss: 0.0390329509973526
step: 500, loss: 0.00024654995650053024
step: 510, loss: 0.07213616371154785
step: 520, loss: 7.020636985544115e-05
step: 530, loss: 0.0020632778760045767
/usr/local/lib/python3.7/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=0.9504483246814535, f1=0.9382948657560056,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010506818071007729
step: 10, loss: 0.017617179080843925
step: 20, loss: 6.439274875447154e-05
step: 30, loss: 0.10992991924285889
step: 40, loss: 0.0001306618651142344
step: 50, loss: 0.0029243130702525377
step: 60, loss: 0.0006657204357907176
step: 70, loss: 0.011630827561020851
step: 80, loss: 0.14406223595142365
step: 90, loss: 0.08816912770271301
step: 100, loss: 0.02699378691613674
step: 110, loss: 0.015396914444863796
step: 120, loss: 0.004446462728083134
step: 130, loss: 0.0021289587020874023
step: 140, loss: 0.005326205864548683
step: 150, loss: 0.017006540670990944
step: 160, loss: 0.00010174342605751008
step: 170, loss: 0.030870074406266212
step: 180, loss: 0.005274270195513964
step: 190, loss: 0.0868438109755516
step: 200, loss: 0.0003013278474099934
step: 210, loss: 0.007664036937057972
step: 220, loss: 0.001111867604777217
step: 230, loss: 8.694881398696452e-05
step: 240, loss: 0.0008613081299699843
step: 250, loss: 0.003942043054848909
step: 260, loss: 0.008931051008403301
step: 270, loss: 0.02446548268198967
step: 280, loss: 0.046392448246479034
step: 290, loss: 0.014838985167443752

```

```

step: 300, loss: 0.0014905165880918503
step: 310, loss: 7.814223499735817e-05
step: 320, loss: 0.0004508514248300344
step: 330, loss: 0.0004189646861050278
step: 340, loss: 0.004906953778117895
step: 350, loss: 0.09036457538604736
step: 360, loss: 0.016468243673443794
step: 370, loss: 0.00020395117462612689
step: 380, loss: 0.00037066690856590867
step: 390, loss: 0.0005245094071142375
step: 400, loss: 0.02067730389535427
step: 410, loss: 0.0055520725436508656
step: 420, loss: 0.0019243924180045724
step: 430, loss: 0.0058988844975829124
step: 440, loss: 0.00034722709096968174
step: 450, loss: 0.01746545359492302
step: 460, loss: 0.0005827692802995443
step: 470, loss: 0.003870014799758792
step: 480, loss: 0.0008064809953793883
step: 490, loss: 0.006963293068110943
step: 500, loss: 0.0287218876183033
step: 510, loss: 0.006326075177639723
step: 520, loss: 0.005857483949512243
step: 530, loss: 0.04311005398631096
/usr/local/lib/python3.7/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=0.9545245751033532, f1=0.9495875343721356,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016946359537541866
step: 10, loss: 0.0017057559452950954
step: 20, loss: 0.0014024346601217985
step: 30, loss: 0.017032641917467117
step: 40, loss: 0.0009731698082759976
step: 50, loss: 0.001159351086243987
step: 60, loss: 0.003887488739565015
step: 70, loss: 0.08671070635318756
step: 80, loss: 0.0019330764189362526
step: 90, loss: 0.0009724590345285833
step: 100, loss: 0.01807895116508007
step: 110, loss: 0.004669359419494867
step: 120, loss: 0.0013821267057210207
step: 130, loss: 0.0013684633886441588
step: 140, loss: 0.00014163237938191742
step: 150, loss: 0.001723324996419251

```

```

step: 160, loss: 0.041350629180669785
step: 170, loss: 8.237065776484087e-05
step: 180, loss: 0.0020701305475085974
step: 190, loss: 2.7281324946670793e-05
step: 200, loss: 0.0001997715880861506
step: 210, loss: 0.06110086292028427
step: 220, loss: 0.00021797452063765377
step: 230, loss: 5.16592881467659e-05
step: 240, loss: 0.0010353625984862447
step: 250, loss: 0.00036983343306928873
step: 260, loss: 0.009486641734838486
step: 270, loss: 0.0005825150292366743
step: 280, loss: 0.0175316222012043
step: 290, loss: 0.0015834912192076445
step: 300, loss: 0.0024628289975225925
step: 310, loss: 0.046130191534757614
step: 320, loss: 0.0011638612486422062
step: 330, loss: 0.05777787044644356
step: 340, loss: 0.00021939298312645406
step: 350, loss: 0.0009972852421924472
step: 360, loss: 0.0011054225033149123
step: 370, loss: 0.0029933154582977295
step: 380, loss: 0.002164258621633053
step: 390, loss: 9.572020644554868e-05
step: 400, loss: 5.7388460845686495e-05
step: 410, loss: 0.0027038827538490295
step: 420, loss: 1.7083555576391518e-05
step: 430, loss: 9.64282444329001e-05
step: 440, loss: 1.2390187293931376e-05
step: 450, loss: 0.0018221024656668305
step: 460, loss: 0.0025295261293649673
step: 470, loss: 0.0029523870907723904
step: 480, loss: 0.022185038775205612
step: 490, loss: 0.0003402139700483531
step: 500, loss: 0.003710906021296978
step: 510, loss: 4.3384701712056994e-05
step: 520, loss: 8.107975008897483e-05
step: 530, loss: 0.0045730313286185265
/usr/local/lib/python3.7/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=0.9542124542124543, f1=0.9499770536943551,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.89962484507123e-06
step: 10, loss: 0.0009185949456878006

```

step: 20, loss: 0.00045994031825102866
step: 30, loss: 2.2126985641079955e-05
step: 40, loss: 0.00022572866873815656
step: 50, loss: 3.9451191696571186e-05
step: 60, loss: 0.0035459038335829973
step: 70, loss: 2.3590157070429996e-05
step: 80, loss: 0.00010210357140749693
step: 90, loss: 0.0011212137760594487
step: 100, loss: 0.0007791469106450677
step: 110, loss: 0.0004346510686445981
step: 120, loss: 0.001710981596261263
step: 130, loss: 0.0001262253208551556
step: 140, loss: 0.0004989180597476661
step: 150, loss: 0.0011105198645964265
step: 160, loss: 0.0006978526362217963
step: 170, loss: 0.0007705076714046299
step: 180, loss: 0.00022345500474330038
step: 190, loss: 0.02360757812857628
step: 200, loss: 0.0001460422354284674
step: 210, loss: 0.00042558921268209815
step: 220, loss: 0.02124403975903988
step: 230, loss: 0.00010902770736720413
step: 240, loss: 0.002124754711985588
step: 250, loss: 0.00018175186414737254
step: 260, loss: 0.0021407685708254576
step: 270, loss: 0.0012380947591736913
step: 280, loss: 0.003300590207800269
step: 290, loss: 0.00011981392162851989
step: 300, loss: 0.04781793802976608
step: 310, loss: 0.020101264119148254
step: 320, loss: 0.001518232747912407
step: 330, loss: 5.674490239471197e-05
step: 340, loss: 0.041359879076480865
step: 350, loss: 9.421911090612411e-05
step: 360, loss: 0.0005362338852137327
step: 370, loss: 0.002391166752204299
step: 380, loss: 0.0014423525426536798
step: 390, loss: 0.0017169601051136851
step: 400, loss: 0.00020402386144269258
step: 410, loss: 0.0007027257815934718
step: 420, loss: 0.014688081108033657
step: 430, loss: 0.0006602771463803947
step: 440, loss: 0.00016147756832651794
step: 450, loss: 0.0014648475917056203
step: 460, loss: 0.0014535801019519567
step: 470, loss: 0.0005931239575147629
step: 480, loss: 0.0004914425662718713
step: 490, loss: 0.0052736313082277775


```

step: 500, loss: 0.029410790652036667
step: 510, loss: 0.07951465249061584
step: 520, loss: 0.0003328395541757345
step: 530, loss: 0.02865997701883316
/usr/local/lib/python3.7/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=0.9471750114836931, f1=0.945337620578778,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035441387444734573
step: 10, loss: 0.0007485010428354144
step: 20, loss: 0.012447117827832699
step: 30, loss: 4.144976264797151e-05
step: 40, loss: 0.002370883710682392
step: 50, loss: 0.004878743551671505
step: 60, loss: 5.6368826335528865e-05
step: 70, loss: 0.35287362337112427
step: 80, loss: 0.0018378666136413813
step: 90, loss: 0.005428516771644354
step: 100, loss: 0.009761886671185493
step: 110, loss: 0.04436719790101051
step: 120, loss: 0.0018705320544540882
step: 130, loss: 0.0008050631731748581
step: 140, loss: 0.0020498023368418217
step: 150, loss: 0.0007764116744510829
step: 160, loss: 0.07289843261241913
step: 170, loss: 0.00012582255294546485
step: 180, loss: 4.232391802361235e-05
step: 190, loss: 0.001502840663306415
step: 200, loss: 0.0011968521866947412
step: 210, loss: 0.00026346519007347524
step: 220, loss: 0.0002020439424086362
step: 230, loss: 6.241998926270753e-05
step: 240, loss: 0.008303877897560596
step: 250, loss: 1.012522534438176e-05
step: 260, loss: 0.0002673354174476117
step: 270, loss: 0.016556378453969955
step: 280, loss: 4.43543067376595e-05
step: 290, loss: 0.0012287698918953538
step: 300, loss: 0.019862012937664986
step: 310, loss: 0.003337386529892683
step: 320, loss: 0.01456140261143446
step: 330, loss: 0.00235663796775043
step: 340, loss: 5.6447821407346055e-05
step: 350, loss: 2.058668360405136e-05

```

```

step: 360, loss: 0.0070128352381289005
step: 370, loss: 0.007248914800584316
step: 380, loss: 7.509417628170922e-05
step: 390, loss: 0.016287218779325485
step: 400, loss: 0.00028374575776979327
step: 410, loss: 0.013818183913826942
step: 420, loss: 0.01296201441437006
step: 430, loss: 0.0016987919807434082
step: 440, loss: 0.0025158762000501156
step: 450, loss: 0.0032338302116841078
step: 460, loss: 0.0010550383012741804
step: 470, loss: 0.0014266886282712221
step: 480, loss: 0.003040294162929058
step: 490, loss: 0.0057455734349787235
step: 500, loss: 0.00011799191997852176
step: 510, loss: 0.012557480484247208
step: 520, loss: 0.0007902121869847178
step: 530, loss: 4.5733067963737994e-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=0.9471264367816092, f1=0.9472234970169803,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002700581680983305
step: 10, loss: 0.0011831398587673903
step: 20, loss: 0.0002510722551960498
step: 30, loss: 2.996525654452853e-05
step: 40, loss: 0.03941088914871216
step: 50, loss: 0.009593834169209003
step: 60, loss: 0.0001283851743210107
step: 70, loss: 0.0011954625369980931
step: 80, loss: 0.0001837392192101106
step: 90, loss: 3.167688919347711e-05
step: 100, loss: 0.0002895394281949848
step: 110, loss: 0.0001571847387822345
step: 120, loss: 4.25617654400412e-05
step: 130, loss: 2.1020654457970522e-05
step: 140, loss: 0.00025662241387180984
step: 150, loss: 0.0020651202648878098
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.0005236220313236117
step: 170, loss: 0.05606778711080551
step: 180, loss: 3.767140515265055e-05
step: 190, loss: 0.000208091878448613
step: 200, loss: 0.07158249616622925

```

```

step: 210, loss: 9.935160960594658e-06
step: 220, loss: 0.02701457403600216
step: 230, loss: 0.0014022699324414134
step: 240, loss: 0.007648753002285957
step: 250, loss: 0.0005752848810516298
step: 260, loss: 6.612471770495176e-05
step: 270, loss: 0.0004239870759192854
step: 280, loss: 0.0003844038874376565
step: 290, loss: 0.0005524393054656684
step: 300, loss: 0.001491791452281177
step: 310, loss: 0.00025440595345571637
step: 320, loss: 7.688505138503388e-05
step: 330, loss: 0.00019777730631176382
step: 340, loss: 0.0028473425190895796
step: 350, loss: 0.00027869694167748094
step: 360, loss: 0.11966507881879807
step: 370, loss: 0.002777791116386652
step: 380, loss: 0.00018952890241052955
step: 390, loss: 9.758664236869663e-05
step: 400, loss: 0.0003400110872462392
step: 410, loss: 7.943552191136405e-05
step: 420, loss: 0.00017188338097184896
step: 430, loss: 9.511393727734685e-05
step: 440, loss: 9.24606774788117e-06
step: 450, loss: 0.018942395225167274
step: 460, loss: 0.00035707856295630336
step: 470, loss: 0.002946528373286128
step: 480, loss: 4.813598934561014e-05
step: 490, loss: 0.0006054331315681338
step: 500, loss: 0.00442213099449873
step: 510, loss: 9.400249109603465e-05
step: 520, loss: 7.710880163358524e-05
step: 530, loss: 1.3824195775669068e-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=0.9475620975160993, f1=0.9418764302059497,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006866732146590948
step: 10, loss: 0.003470565425232053
step: 20, loss: 6.022413072059862e-05
step: 30, loss: 0.0020772006828337908
step: 40, loss: 0.0002707380335777998
step: 50, loss: 0.021110380068421364
step: 60, loss: 2.1209996702964418e-05

```

step: 70, loss: 7.451279816450551e-05
step: 80, loss: 9.653118468122557e-05
step: 90, loss: 1.180129675049102e-05
step: 100, loss: 5.3943604143569246e-05
step: 110, loss: 0.0015095015987753868
step: 120, loss: 6.2994386098580435e-06
step: 130, loss: 0.00013630879402626306
step: 140, loss: 0.0009778719395399094
step: 150, loss: 0.0008207603823393583
step: 160, loss: 0.0005173300742171705
step: 170, loss: 0.004442967474460602
step: 180, loss: 0.00017640912847127765
step: 190, loss: 0.0002251030964544043
step: 200, loss: 6.615442543989047e-05
step: 210, loss: 0.002037404803559184
step: 220, loss: 3.786753222811967e-05
step: 230, loss: 0.0034977111499756575
step: 240, loss: 0.0012094553094357252
step: 250, loss: 0.004208695143461227
step: 260, loss: 7.816607103450224e-05
step: 270, loss: 0.000686628045514226
step: 280, loss: 5.005057028029114e-05
step: 290, loss: 1.1183155947946943e-05
step: 300, loss: 1.0937181286863051e-05
step: 310, loss: 2.0454539480851963e-05
step: 320, loss: 0.00048262806376442313
step: 330, loss: 0.0006260026129893959
step: 340, loss: 0.00022857898147776723
step: 350, loss: 0.0008247795049101114
step: 360, loss: 4.775681009050459e-05
step: 370, loss: 0.0014453394105657935
step: 380, loss: 0.06975223124027252
step: 390, loss: 0.00025145182735286653
step: 400, loss: 0.0015411016065627337
step: 410, loss: 3.907893187715672e-05
step: 420, loss: 0.02056911215186119
step: 430, loss: 0.00046847155317664146
step: 440, loss: 0.003745677648112178
step: 450, loss: 2.287486677232664e-05
step: 460, loss: 0.025318648666143417
step: 470, loss: 1.368543780699838e-05
step: 480, loss: 0.00015277357306331396
step: 490, loss: 6.562491034856066e-05
step: 500, loss: 0.0012889702338725328
step: 510, loss: 6.419308192562312e-05
step: 520, loss: 0.0008068914175964892
step: 530, loss: 0.0019609297160059214
/usr/local/lib/python3.7/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=0.9522058823529412, f1=0.9431714023831348,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00185682054143399
step: 10, loss: 0.005022736266255379
step: 20, loss: 0.0016193331684917212
step: 30, loss: 0.01649380847811699
step: 40, loss: 1.80962397280382e-05
step: 50, loss: 0.0006589906406588852
step: 60, loss: 0.0016209259629249573
step: 70, loss: 1.6360516383429058e-05
step: 80, loss: 0.003048927756026387
step: 90, loss: 0.00017082194972317666
step: 100, loss: 7.835992437321693e-05
step: 110, loss: 0.003931701183319092
step: 120, loss: 8.78464343259111e-05
step: 130, loss: 0.0025289750192314386
step: 140, loss: 0.0011027700966224074
step: 150, loss: 6.0444963310146704e-05
step: 160, loss: 0.0009118210291489959
step: 170, loss: 1.4788932276132982e-05
step: 180, loss: 2.6644520403351635e-05
step: 190, loss: 0.003125068498775363
step: 200, loss: 0.0016613946063444018
step: 210, loss: 0.0014104750007390976
step: 220, loss: 1.767134563124273e-05
step: 230, loss: 0.001040104078128934
step: 240, loss: 2.0569188563968055e-05
step: 250, loss: 7.399357855319977e-05
step: 260, loss: 1.3816598766425159e-05
step: 270, loss: 5.7274468417745084e-05
step: 280, loss: 2.1899757484789006e-05
step: 290, loss: 0.0004913643933832645
step: 300, loss: 4.198173337499611e-05
step: 310, loss: 0.0453202985227108
step: 320, loss: 4.199319664621726e-05
step: 330, loss: 1.7228348951903172e-05
step: 340, loss: 1.968674587260466e-05
step: 350, loss: 0.0018546809442341328
step: 360, loss: 0.0005922937416471541
step: 370, loss: 1.4613352504966315e-05
step: 380, loss: 0.0009397954563610256
step: 390, loss: 0.003909976687282324
step: 400, loss: 0.005777440033853054

```

```

step: 410, loss: 7.805398490745574e-05
step: 420, loss: 0.0022979455534368753
step: 430, loss: 0.00040797522524371743
step: 440, loss: 0.003947754390537739
step: 450, loss: 0.004116402473300695
step: 460, loss: 0.017666101455688477
step: 470, loss: 3.478955113678239e-05
step: 480, loss: 0.0034676562063395977
step: 490, loss: 0.00022699471446685493
step: 500, loss: 0.0024785404093563557
step: 510, loss: 3.542509148246609e-05
step: 520, loss: 1.636038905417081e-05
step: 530, loss: 1.3753485291090328e-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=0.9513314967860421, f1=0.945537757437071,
best_f1=0.9489939167056621
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 [00:23, 241.54it/s]
load_f1 = 0.9508804448563485
real_f1 = 0.9497464269248502
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 192.70it/s]

```

4.1.4 Amazon-Google - Running the matcher

```

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

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

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

```

Selected optimization level 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: 0.5204543471336365
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44445592164993286
step: 20, loss: 0.48205116391181946
step: 30, loss: 0.346344530582428
step: 40, loss: 0.32370299100875854
step: 50, loss: 0.5296871662139893
step: 60, loss: 0.40523603558540344
step: 70, loss: 0.31843239068984985
step: 80, loss: 0.3878343999385834
step: 90, loss: 0.25449255108833313
step: 100, loss: 0.2612176835536957
step: 110, loss: 0.2862390875816345
step: 120, loss: 0.3676747977733612

```



```

step: 130, loss: 0.23270460963249207
step: 140, loss: 0.49607983231544495
step: 150, loss: 0.30957165360450745
step: 160, loss: 0.346572607755661
step: 170, loss: 0.19839709997177124
step: 180, loss: 0.271323025226593
step: 190, loss: 0.48896607756614685
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 200, loss: 0.32567888498306274
step: 210, loss: 0.34461069107055664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5504587155963303, f1=0.5571142284569138,
best_f1=0.5571142284569138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25362294912338257
step: 10, loss: 0.0943836122751236
step: 20, loss: 0.31294354796409607
step: 30, loss: 0.3268393874168396
step: 40, loss: 0.4135677218437195
step: 50, loss: 0.21058088541030884
step: 60, loss: 0.3648391366004944
step: 70, loss: 0.19315780699253082
step: 80, loss: 0.19536042213439941
step: 90, loss: 0.14596740901470184
step: 100, loss: 0.4408653676509857
step: 110, loss: 0.45024287700653076
step: 120, loss: 0.159770205616951
step: 130, loss: 0.21259279549121857
step: 140, loss: 0.1993192434310913
step: 150, loss: 0.2745380997657776
step: 160, loss: 0.05913499742746353
step: 170, loss: 0.25612786412239075
step: 180, loss: 0.26973822712898254
step: 190, loss: 0.2836220860481262
step: 200, loss: 0.10587811470031738
step: 210, loss: 0.19501711428165436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.6432989690721649, f1=0.6030368763557483,

best_f1=0.6030368763557483

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

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

step: 0, loss: 0.0924774631857872

step: 10, loss: 0.07212181389331818

step: 20, loss: 0.1407814770936966

step: 30, loss: 0.06399515271186829

step: 40, loss: 0.3169785141944885

step: 50, loss: 0.18412628769874573

step: 60, loss: 0.26227933168411255

step: 70, loss: 0.18099839985370636

step: 80, loss: 0.15664520859718323

step: 90, loss: 0.09374655783176422

step: 100, loss: 0.15528488159179688

step: 110, loss: 0.1418510377407074

step: 120, loss: 0.17401796579360962

step: 130, loss: 0.13241161406040192

step: 140, loss: 0.24967026710510254

step: 150, loss: 0.1830185502767563

step: 160, loss: 0.10727623850107193

step: 170, loss: 0.17548535764217377

step: 180, loss: 0.11424857378005981

step: 190, loss: 0.06496866792440414

step: 200, loss: 0.14333534240722656

step: 210, loss: 0.21308235824108124

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

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

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

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=0.6814516129032259, f1=0.6736401673640167,

best_f1=0.6736401673640167

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

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

step: 0, loss: 0.09836465120315552

step: 10, loss: 0.13565264642238617

step: 20, loss: 0.0714925080537796

step: 30, loss: 0.07131947576999664

step: 40, loss: 0.04665521904826164

step: 50, loss: 0.14319421350955963

```

step: 60, loss: 0.26179441809654236
step: 70, loss: 0.0833132341504097
step: 80, loss: 0.0829327255487442
step: 90, loss: 0.06923015415668488
step: 100, loss: 0.13550354540348053
step: 110, loss: 0.23483745753765106
step: 120, loss: 0.19380097091197968
step: 130, loss: 0.33453577756881714
step: 140, loss: 0.13131281733512878
step: 150, loss: 0.09713460505008698
step: 160, loss: 0.20997831225395203
step: 170, loss: 0.09199447184801102
step: 180, loss: 0.03308021277189255
step: 190, loss: 0.1660076081752777
step: 200, loss: 0.03141496330499649
step: 210, loss: 0.4191853702068329
/usr/local/lib/python3.7/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=0.6802218114602587, f1=0.6730038022813688,
best_f1=0.6736401673640167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1431507021188736
step: 10, loss: 0.058938659727573395
step: 20, loss: 0.12665033340454102
step: 30, loss: 0.03537560626864433
step: 40, loss: 0.0766279399394989
step: 50, loss: 0.16420520842075348
step: 60, loss: 0.11880286782979965
step: 70, loss: 0.1627875417470932
step: 80, loss: 0.0922975093126297
step: 90, loss: 0.12003066390752792
step: 100, loss: 0.02341023087501526
step: 110, loss: 0.06789064407348633
step: 120, loss: 0.14829108119010925
step: 130, loss: 0.06113245710730553
step: 140, loss: 0.11391206085681915
step: 150, loss: 0.10930989682674408
step: 160, loss: 0.03712089732289314
step: 170, loss: 0.1775893270969391
step: 180, loss: 0.12807190418243408
step: 190, loss: 0.18507450819015503
step: 200, loss: 0.06378979980945587
step: 210, loss: 0.0908391922712326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7171314741035857, f1=0.6880000000000001,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010551825165748596
step: 10, loss: 0.06120726838707924
step: 20, loss: 0.04073294997215271
step: 30, loss: 0.09850561618804932
step: 40, loss: 0.04573439806699753
step: 50, loss: 0.07860098034143448
step: 60, loss: 0.13739009201526642
step: 70, loss: 0.10862398892641068
step: 80, loss: 0.14141155779361725
step: 90, loss: 0.039288148283958435
step: 100, loss: 0.17580841481685638
step: 110, loss: 0.07208168506622314
step: 120, loss: 0.02343217469751835
step: 130, loss: 0.07656055688858032
step: 140, loss: 0.07869163900613785
step: 150, loss: 0.05555390194058418
step: 160, loss: 0.10045108944177628
step: 170, loss: 0.14116324484348297
step: 180, loss: 0.10796378552913666
step: 190, loss: 0.03591613844037056
step: 200, loss: 0.05397653952240944
step: 210, loss: 0.08426148444414139
/usr/local/lib/python3.7/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=0.6945454545454546, f1=0.700374531835206,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11613983660936356
step: 10, loss: 0.012826166115701199
step: 20, loss: 0.10999773442745209
step: 30, loss: 0.085934579372406
step: 40, loss: 0.10494457185268402
step: 50, loss: 0.11497004330158234
step: 60, loss: 0.04273567721247673
step: 70, loss: 0.011110448278486729

```

```

step: 80, loss: 0.10911287367343903
step: 90, loss: 0.23498210310935974
step: 100, loss: 0.03658336028456688
step: 110, loss: 0.20453180372714996
step: 120, loss: 0.12468200922012329
step: 130, loss: 0.12000086903572083
step: 140, loss: 0.06918969005346298
step: 150, loss: 0.042039643973112106
step: 160, loss: 0.2838110625743866
step: 170, loss: 0.13915051519870758
step: 180, loss: 0.023063089698553085
step: 190, loss: 0.05585508048534393
step: 200, loss: 0.044649310410022736
step: 210, loss: 0.20069970190525055
/usr/local/lib/python3.7/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=0.7065420560747664, f1=0.7047619047619047,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08134378492832184
step: 10, loss: 0.06561563163995743
step: 20, loss: 0.08877004683017731
step: 30, loss: 0.02099023386836052
step: 40, loss: 0.019045019522309303
step: 50, loss: 0.02572118304669857
step: 60, loss: 0.01041332446038723
step: 70, loss: 0.07301641255617142
step: 80, loss: 0.05404755845665932
step: 90, loss: 0.04567752406001091
step: 100, loss: 0.10681506246328354
step: 110, loss: 0.011736146174371243
step: 120, loss: 0.19489572942256927
step: 130, loss: 0.01481566857546568
step: 140, loss: 0.09489526599645615
step: 150, loss: 0.1715685874223709
step: 160, loss: 0.22427624464035034
step: 170, loss: 0.1718294471502304
step: 180, loss: 0.025392621755599976
step: 190, loss: 0.01404124777764082
step: 200, loss: 0.07400860637426376
step: 210, loss: 0.1550140529870987
/usr/local/lib/python3.7/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=0.70703125, f1=0.7103174603174603, best_f1=0.6880000000000001

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01996748521924019
step: 10, loss: 0.1704399585723877
step: 20, loss: 0.030652571469545364
step: 30, loss: 0.016635362058877945
step: 40, loss: 0.14725515246391296
step: 50, loss: 0.021704431623220444
step: 60, loss: 0.01981111615896225
step: 70, loss: 0.11935162544250488
step: 80, loss: 0.008596144616603851
step: 90, loss: 0.003138526575639844
step: 100, loss: 0.04981504753232002
step: 110, loss: 0.10894042253494263
step: 120, loss: 0.1648443192243576
step: 130, loss: 0.04733644053339958
step: 140, loss: 0.03595918416976929
step: 150, loss: 0.023212825879454613
step: 160, loss: 0.14858393371105194
step: 170, loss: 0.08493472635746002
step: 180, loss: 0.04680859297513962
step: 190, loss: 0.02061157301068306
step: 200, loss: 0.03568913787603378
step: 210, loss: 0.028752649202942848
/usr/local/lib/python3.7/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=0.6878504672897197, f1=0.6996197718631179,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005333304405212402
step: 10, loss: 0.009074311703443527
step: 20, loss: 0.0366930291056633
step: 30, loss: 0.005898197181522846
step: 40, loss: 0.09324635565280914
step: 50, loss: 0.05002511665225029
step: 60, loss: 0.010677514597773552
step: 70, loss: 0.06868229806423187
step: 80, loss: 0.01634952612221241
step: 90, loss: 0.05809323862195015
step: 100, loss: 0.06252268701791763
step: 110, loss: 0.08567079156637192
step: 120, loss: 0.05921221151947975
step: 130, loss: 0.02497466281056404
step: 140, loss: 0.048718880861997604

```

```

step: 150, loss: 0.013442067429423332
step: 160, loss: 0.025134166702628136
step: 170, loss: 0.010556574910879135
step: 180, loss: 0.01775442063808441
step: 190, loss: 0.02277514711022377
step: 200, loss: 0.05829544737935066
step: 210, loss: 0.04992671310901642
/usr/local/lib/python3.7/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=0.7056672760511882, f1=0.7026022304832713,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0591551847755909
step: 10, loss: 0.03742740675806999
step: 20, loss: 0.08465297520160675
step: 30, loss: 0.04610735923051834
step: 40, loss: 0.05442475900053978
step: 50, loss: 0.024905338883399963
step: 60, loss: 0.02981467731297016
step: 70, loss: 0.02348119206726551
step: 80, loss: 0.018833590671420097
step: 90, loss: 0.08011326193809509
step: 100, loss: 0.02350062131881714
step: 110, loss: 0.039860133081674576
step: 120, loss: 0.05309680849313736
step: 130, loss: 0.010692841373383999
step: 140, loss: 0.06742221862077713
step: 150, loss: 0.05523126572370529
step: 160, loss: 0.005682777147740126
step: 170, loss: 0.05679449066519737
step: 180, loss: 0.013057312928140163
step: 190, loss: 0.04346107318997383
step: 200, loss: 0.01512033399194479
step: 210, loss: 0.03875814378261566
/usr/local/lib/python3.7/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=0.7111984282907662, f1=0.7004048582995952,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026921352837234735
step: 10, loss: 0.0016754630487412214
step: 20, loss: 0.16084198653697968

```

```

step: 30, loss: 0.0033356046769768
step: 40, loss: 0.003332403488457203
step: 50, loss: 0.009287690743803978
step: 60, loss: 0.004273788072168827
step: 70, loss: 0.03149279206991196
step: 80, loss: 0.12696312367916107
step: 90, loss: 0.17355765402317047
step: 100, loss: 0.0025914390571415424
step: 110, loss: 0.005097212269902229
step: 120, loss: 0.0019363160245120525
step: 130, loss: 0.01866072230041027
step: 140, loss: 0.2781675457954407
step: 150, loss: 0.0037419020663946867
step: 160, loss: 0.1003216877579689
step: 170, loss: 0.041295379400253296
step: 180, loss: 0.06763063371181488
step: 190, loss: 0.13111580908298492
step: 200, loss: 0.018630243837833405
step: 210, loss: 0.06293804198503494
/usr/local/lib/python3.7/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=0.7047970479704797, f1=0.7085714285714286,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04315698891878128
step: 10, loss: 0.04929037392139435
step: 20, loss: 0.010230854153633118
step: 30, loss: 0.003603088203817606
step: 40, loss: 0.013619872741401196
step: 50, loss: 0.07097884267568588
step: 60, loss: 0.0037598346825689077
step: 70, loss: 0.0025797730777412653
step: 80, loss: 0.03346508741378784
step: 90, loss: 0.006857460364699364
step: 100, loss: 0.020555617287755013
step: 110, loss: 0.12637004256248474
step: 120, loss: 0.012831811793148518
step: 130, loss: 0.002184899291023612
step: 140, loss: 0.05193113908171654
step: 150, loss: 0.002950756112113595
step: 160, loss: 0.06999637931585312
step: 170, loss: 0.08585961908102036
step: 180, loss: 0.09687408804893494
step: 190, loss: 0.02572697028517723
step: 200, loss: 0.06374017149209976

```



```

step: 210, loss: 0.040550801903009415
/usr/local/lib/python3.7/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=0.7100591715976332, f1=0.7030303030303031,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009266631677746773
step: 10, loss: 0.04212541878223419
step: 20, loss: 0.007899587973952293
step: 30, loss: 0.0032919312361627817
step: 40, loss: 0.15613529086112976
step: 50, loss: 0.021454960107803345
step: 60, loss: 0.03840342164039612
step: 70, loss: 0.03520498424768448
step: 80, loss: 0.022338127717375755
step: 90, loss: 0.0074179088696837425
step: 100, loss: 0.0058292546309530735
step: 110, loss: 0.003142670262604952
step: 120, loss: 0.006605623289942741
step: 130, loss: 0.020570239052176476
step: 140, loss: 0.00830624345690012
step: 150, loss: 0.13285087049007416
step: 160, loss: 0.0045492565259337425
step: 170, loss: 0.020417578518390656
step: 180, loss: 0.001996460137888789
step: 190, loss: 0.050434403121471405
step: 200, loss: 0.02524098940193653
step: 210, loss: 0.08365632593631744
/usr/local/lib/python3.7/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=0.700990099009901, f1=0.6987951807228915,
best_f1=0.6880000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011488246731460094
step: 10, loss: 0.00917941052466631
step: 20, loss: 0.03725787624716759
step: 30, loss: 0.059596024453639984
step: 40, loss: 0.03672037646174431
step: 50, loss: 0.002678225515410304
step: 60, loss: 0.008130096830427647
step: 70, loss: 0.04242724925279617
step: 80, loss: 0.10052026808261871

```

```

step: 90, loss: 0.04838049039244652
step: 100, loss: 0.018780097365379333
step: 110, loss: 0.09745516628026962
step: 120, loss: 0.04439961910247803
step: 130, loss: 0.16997268795967102
step: 140, loss: 0.014169322326779366
step: 150, loss: 0.07656436413526535
step: 160, loss: 0.004000192973762751
step: 170, loss: 0.004483371507376432
step: 180, loss: 0.009357243776321411
step: 190, loss: 0.0034678014926612377
step: 200, loss: 0.0022190159652382135
step: 210, loss: 0.03912757337093353
/usr/local/lib/python3.7/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=0.7140186915887851, f1=0.7049808429118773,
best_f1=0.6880000000000001
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

2293it [00:05, 448.06it/s]
load_f1 = 0.7121771217712177
real_f1 = 0.7096774193548387
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:23, 189.53it/s]

```

4.1.5 Walmart-Amazon - Running the matcher

```

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

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

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

```

Selected optimization level 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: 0.46809953451156616
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44124162197113037
step: 20, loss: 0.273418128490448
step: 30, loss: 0.4198245406150818
step: 40, loss: 0.2516299784183502
step: 50, loss: 0.3079386353492737
step: 60, loss: 0.44406652450561523
step: 70, loss: 0.45767831802368164
step: 80, loss: 0.1470637023448944
step: 90, loss: 0.3610319197177887
step: 100, loss: 0.5344370603561401
step: 110, loss: 0.2369273453950882
step: 120, loss: 0.3170967400074005

```

```

step: 130, loss: 0.2993752658367157
step: 140, loss: 0.16079939901828766
step: 150, loss: 0.3409379720687866
step: 160, loss: 0.24357371032238007
step: 170, loss: 0.3721579611301422
step: 180, loss: 0.16750314831733704
step: 190, loss: 0.1566644012928009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3929799497127533
step: 10, loss: 0.32362979650497437
step: 20, loss: 0.6283731460571289
step: 30, loss: 0.2544998824596405
step: 40, loss: 0.5676438212394714
step: 50, loss: 0.31128641963005066
step: 60, loss: 0.46546876430511475
step: 70, loss: 0.32100218534469604
step: 80, loss: 0.15285469591617584
step: 90, loss: 0.31726256012916565
step: 100, loss: 0.24430140852928162
step: 110, loss: 0.37549006938934326
step: 120, loss: 0.22873279452323914
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5647369623184204
step: 140, loss: 0.3069729506969452
step: 150, loss: 0.3113163411617279
step: 160, loss: 0.31635582447052
step: 170, loss: 0.24171464145183563
step: 180, loss: 0.17084750533103943
step: 190, loss: 0.24224776029586792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.17255252570406795, f1=0.170993733213966,

```

```

best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3773976266384125
step: 10, loss: 0.38875091075897217
step: 20, loss: 0.46542078256607056
step: 30, loss: 0.32381564378738403
step: 40, loss: 0.08750315755605698
step: 50, loss: 0.38859158754348755
step: 60, loss: 0.16112835705280304
step: 70, loss: 0.3784819543361664
step: 80, loss: 0.3132477104663849
step: 90, loss: 0.36997848749160767
step: 100, loss: 0.5579948425292969
step: 110, loss: 0.6594672203063965
step: 120, loss: 0.3712177574634552
step: 130, loss: 0.16503696143627167
step: 140, loss: 0.37897545099258423
step: 150, loss: 0.3035449981689453
step: 160, loss: 0.6295177340507507
step: 170, loss: 0.4481334984302521
step: 180, loss: 0.3913755714893341
step: 190, loss: 0.16774539649486542
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24049603939056396
step: 10, loss: 0.24121670424938202
step: 20, loss: 0.3114635944366455
step: 30, loss: 0.23059910535812378
step: 40, loss: 0.5472692847251892
step: 50, loss: 0.2522914409637451
step: 60, loss: 0.37973201274871826
step: 70, loss: 0.3202034831047058
step: 80, loss: 0.2333616316318512
step: 90, loss: 0.17413434386253357
step: 100, loss: 0.32002705335617065
step: 110, loss: 0.369820237159729
step: 120, loss: 0.2584103047847748
step: 130, loss: 0.48201441764831543
step: 140, loss: 0.29965949058532715
step: 150, loss: 0.23375371098518372

```

```

step: 160, loss: 0.3042123019695282
step: 170, loss: 0.4506773054599762
step: 180, loss: 0.37391528487205505
step: 190, loss: 0.16411082446575165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3877078592777252
step: 10, loss: 0.38815534114837646
step: 20, loss: 0.16553238034248352
step: 30, loss: 0.11227987706661224
step: 40, loss: 0.3122049570083618
step: 50, loss: 0.5132584571838379
step: 60, loss: 0.24661991000175476
step: 70, loss: 0.38795197010040283
step: 80, loss: 0.3730844259262085
step: 90, loss: 0.31018954515457153
step: 100, loss: 0.4553123116493225
step: 110, loss: 0.406453937292099
step: 120, loss: 0.23014920949935913
step: 130, loss: 0.5578641295433044
step: 140, loss: 0.3854200839996338
step: 150, loss: 0.31442368030548096
step: 160, loss: 0.16527828574180603
step: 170, loss: 0.38708725571632385
step: 180, loss: 0.24737673997879028
step: 190, loss: 0.31603190302848816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31236815452575684
step: 10, loss: 0.25195637345314026
step: 20, loss: 0.3163641393184662
step: 30, loss: 0.47537392377853394
step: 40, loss: 0.24648517370224
step: 50, loss: 0.31295669078826904
step: 60, loss: 0.4460686445236206
step: 70, loss: 0.311991423368454

```

```

step: 80, loss: 0.31610459089279175
step: 90, loss: 0.23177161812782288
step: 100, loss: 0.49651023745536804
step: 110, loss: 0.23813335597515106
step: 120, loss: 0.4588446021080017
step: 130, loss: 0.5209100246429443
step: 140, loss: 0.18974804878234863
step: 150, loss: 0.383613646030426
step: 160, loss: 0.3817117214202881
step: 170, loss: 0.3801420032978058
step: 180, loss: 0.17798203229904175
step: 190, loss: 0.308363139629364
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3067079186439514
step: 10, loss: 0.31778815388679504
step: 20, loss: 0.30850768089294434
step: 30, loss: 0.18696622550487518
step: 40, loss: 0.31940826773643494
step: 50, loss: 0.08845797926187515
step: 60, loss: 0.15647682547569275
step: 70, loss: 0.15838924050331116
step: 80, loss: 0.2319052666425705
step: 90, loss: 0.2340557724237442
step: 100, loss: 0.5710362792015076
step: 110, loss: 0.4456811845302582
step: 120, loss: 0.38255414366722107
step: 130, loss: 0.3157128095626831
step: 140, loss: 0.24717092514038086
step: 150, loss: 0.3109336197376251
step: 160, loss: 0.3765864074230194
step: 170, loss: 0.3126183748245239
step: 180, loss: 0.2377414107322693
step: 190, loss: 0.3144899010658264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.23826007544994354
step: 10, loss: 0.3877915143966675
step: 20, loss: 0.24048307538032532
step: 30, loss: 0.3803219795227051
step: 40, loss: 0.16971951723098755
step: 50, loss: 0.31196334958076477
step: 60, loss: 0.45370253920555115
step: 70, loss: 0.24155788123607635
step: 80, loss: 0.388068825006485
step: 90, loss: 0.24134831130504608
step: 100, loss: 0.30892252922058105
step: 110, loss: 0.3784017860889435
step: 120, loss: 0.39314982295036316
step: 130, loss: 0.3185468912124634
step: 140, loss: 0.3863335847854614
step: 150, loss: 0.3188410997390747
step: 160, loss: 0.19720429182052612
step: 170, loss: 0.247055321931839
step: 180, loss: 0.3072035312652588
step: 190, loss: 0.31105902791023254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2446664273738861
step: 10, loss: 0.09769579768180847
step: 20, loss: 0.3125000298023224
step: 30, loss: 0.18309447169303894
step: 40, loss: 0.3139534294605255
step: 50, loss: 0.3795013725757599
step: 60, loss: 0.378933846950531
step: 70, loss: 0.16779421269893646
step: 80, loss: 0.3910621702671051
step: 90, loss: 0.703658401966095
step: 100, loss: 0.3786983788013458
step: 110, loss: 0.37393972277641296
step: 120, loss: 0.586990475654602
step: 130, loss: 0.3106079697608948
step: 140, loss: 0.3123110830783844
step: 150, loss: 0.2389499545097351
step: 160, loss: 0.23198562860488892
step: 170, loss: 0.4603809714317322
step: 180, loss: 0.30456435680389404
step: 190, loss: 0.2424805760383606

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2404877245426178
step: 10, loss: 0.17320232093334198
step: 20, loss: 0.30556198954582214
step: 30, loss: 0.37894442677497864
step: 40, loss: 0.30690234899520874
step: 50, loss: 0.09904744476079941
step: 60, loss: 0.31820398569107056
step: 70, loss: 0.30769187211990356
step: 80, loss: 0.38161465525627136
step: 90, loss: 0.1609327495098114
step: 100, loss: 0.23597395420074463
step: 110, loss: 0.2345932424068451
step: 120, loss: 0.38448283076286316
step: 130, loss: 0.5185496211051941
step: 140, loss: 0.38595110177993774
step: 150, loss: 0.3057057559490204
step: 160, loss: 0.253286749124527
step: 170, loss: 0.3796997666358948
step: 180, loss: 0.23863625526428223
step: 190, loss: 0.16176390647888184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5385518074035645
step: 10, loss: 0.09501645714044571
step: 20, loss: 0.23448780179023743
step: 30, loss: 0.23736734688282013
step: 40, loss: 0.08939812332391739
step: 50, loss: 0.31512919068336487
step: 60, loss: 0.23743148148059845
step: 70, loss: 0.6232336759567261
step: 80, loss: 0.37509679794311523
step: 90, loss: 0.3077530860900879
step: 100, loss: 0.17253656685352325
step: 110, loss: 0.2455463856458664

```

```

step: 120, loss: 0.30644652247428894
step: 130, loss: 0.5951719284057617
step: 140, loss: 0.4648991823196411
step: 150, loss: 0.31164437532424927
step: 160, loss: 0.3109680116176605
step: 170, loss: 0.4427157938480377
step: 180, loss: 0.314277321100235
step: 190, loss: 0.11206144094467163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39014461636543274
step: 10, loss: 0.3094594478607178
step: 20, loss: 0.5326707363128662
step: 30, loss: 0.37707364559173584
step: 40, loss: 0.30448704957962036
step: 50, loss: 0.5205159783363342
step: 60, loss: 0.4611339569091797
step: 70, loss: 0.38837218284606934
step: 80, loss: 0.38481757044792175
step: 90, loss: 0.4526844322681427
step: 100, loss: 0.3137587904930115
step: 110, loss: 0.45501843094825745
step: 120, loss: 0.17064246535301208
step: 130, loss: 0.23620331287384033
step: 140, loss: 0.3741360604763031
step: 150, loss: 0.37340325117111206
step: 160, loss: 0.24377384781837463
step: 170, loss: 0.44862255454063416
step: 180, loss: 0.3115696609020233
step: 190, loss: 0.1715843230485916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3063986003398895
step: 10, loss: 0.30692270398139954
step: 20, loss: 0.3862365782260895
step: 30, loss: 0.38536185026168823

```

```

step: 40, loss: 0.24732144176959991
step: 50, loss: 0.2413519024848938
step: 60, loss: 0.3173897862434387
step: 70, loss: 0.16327396035194397
step: 80, loss: 0.3126620352268219
step: 90, loss: 0.3066665828227997
step: 100, loss: 0.16725915670394897
step: 110, loss: 0.5976550579071045
step: 120, loss: 0.16718338429927826
step: 130, loss: 0.23875540494918823
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.30774369835853577
step: 150, loss: 0.3760443329811096
step: 160, loss: 0.45520535111427307
step: 170, loss: 0.10299882292747498
step: 180, loss: 0.23936845362186432
step: 190, loss: 0.3117547035217285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31128832697868347
step: 10, loss: 0.2425956428050995
step: 20, loss: 0.39247727394104004
step: 30, loss: 0.10381759703159332
step: 40, loss: 0.3122803866863251
step: 50, loss: 0.38104215264320374
step: 60, loss: 0.3156180679798126
step: 70, loss: 0.2400793582201004
step: 80, loss: 0.39648333191871643
step: 90, loss: 0.165862575173378
step: 100, loss: 0.3145937919616699
step: 110, loss: 0.45210933685302734
step: 120, loss: 0.30896613001823425
step: 130, loss: 0.31256037950515747
step: 140, loss: 0.17226073145866394
step: 150, loss: 0.455392986536026
step: 160, loss: 0.3813740909099579
step: 170, loss: 0.23452100157737732
step: 180, loss: 0.30747637152671814
step: 190, loss: 0.5263765454292297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

epoch 14: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5374159216880798
step: 10, loss: 0.24098488688468933
step: 20, loss: 0.24115359783172607
step: 30, loss: 0.3068667948246002
step: 40, loss: 0.654503583908081
step: 50, loss: 0.31562161445617676
step: 60, loss: 0.17333218455314636
step: 70, loss: 0.25012606382369995
step: 80, loss: 0.23803509771823883
step: 90, loss: 0.31348690390586853
step: 100, loss: 0.3791303336620331
step: 110, loss: 0.4522913098335266
step: 120, loss: 0.235283762216568
step: 130, loss: 0.3042800724506378
step: 140, loss: 0.4519476890563965
step: 150, loss: 0.1689731329679489
step: 160, loss: 0.3120405972003937
step: 170, loss: 0.7463729381561279
step: 180, loss: 0.30890387296676636
step: 190, loss: 0.45753175020217896
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.170993733213966
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, 238.18it/s]
load_f1 = 0.1724137931034483
real_f1 = 0.17239839213934793
733it [00:00, 3462.58it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 189.79it/s]

```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

```

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

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

```

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.4796449542045593
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.49684906005859375
step: 20, loss: 0.2734176814556122
step: 30, loss: 0.3796963095664978
step: 40, loss: 0.5247427225112915
step: 50, loss: 0.3416529595851898
step: 60, loss: 0.5236034393310547
step: 70, loss: 0.3107782006263733
step: 80, loss: 0.25894659757614136
step: 90, loss: 0.2133503556251526
step: 100, loss: 0.18393012881278992
step: 110, loss: 0.38099223375320435
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.29460930824279785
step: 130, loss: 0.3079601228237152
step: 140, loss: 0.38826894760131836
step: 150, loss: 0.30827468633651733
step: 160, loss: 0.4028743505477905
step: 170, loss: 0.3258276879787445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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=0.19615384615384615, f1=0.19509851033157133,
best_f1=0.19509851033157133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31310898065567017
step: 10, loss: 0.4544471800327301
step: 20, loss: 0.3106052577495575
step: 30, loss: 0.30191949009895325
step: 40, loss: 0.05828477814793587
step: 50, loss: 0.46576541662216187
step: 60, loss: 0.1726066917181015
step: 70, loss: 0.49339327216148376
step: 80, loss: 0.22759319841861725
step: 90, loss: 0.24887855350971222
step: 100, loss: 0.5239108800888062
step: 110, loss: 0.2609930634498596
step: 120, loss: 0.2569308876991272
step: 130, loss: 0.541729211807251
step: 140, loss: 0.5253250598907471

```



```

step: 150, loss: 0.41496938467025757
step: 160, loss: 0.4195663332939148
step: 170, loss: 0.4122472107410431
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.19509851033157133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6104037761688232
step: 10, loss: 0.32911673188209534
step: 20, loss: 0.2405547797679901
step: 30, loss: 0.25820404291152954
step: 40, loss: 0.3714671730995178
step: 50, loss: 0.5883455872535706
step: 60, loss: 0.3161734342575073
step: 70, loss: 0.23052072525024414
step: 80, loss: 0.3719688951969147
step: 90, loss: 0.5422374606132507
step: 100, loss: 0.33540067076683044
step: 110, loss: 0.16701389849185944
step: 120, loss: 0.5885420441627502
step: 130, loss: 0.4839666485786438
step: 140, loss: 0.46025019884109497
step: 150, loss: 0.19927649199962616
step: 160, loss: 0.17334763705730438
step: 170, loss: 0.31138256192207336
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.19509851033157133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3754129707813263
step: 10, loss: 0.5372826457023621
step: 20, loss: 0.2506229877471924
step: 30, loss: 0.43891987204551697
step: 40, loss: 0.24654383957386017
step: 50, loss: 0.3852066099643707
step: 60, loss: 0.7068683505058289
step: 70, loss: 0.3425576388835907
step: 80, loss: 0.4391467869281769
step: 90, loss: 0.3069654703140259
step: 100, loss: 0.3746148943901062

```

```

step: 110, loss: 0.4431656301021576
step: 120, loss: 0.4898681044578552
step: 130, loss: 0.3033302426338196
step: 140, loss: 0.2388259321451187
step: 150, loss: 0.6722103953361511
step: 160, loss: 0.12751434743404388
step: 170, loss: 0.24266043305397034
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.19509851033157133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37829726934432983
step: 10, loss: 0.3253650367259979
step: 20, loss: 0.318024218082428
step: 30, loss: 0.30747196078300476
step: 40, loss: 0.25065314769744873
step: 50, loss: 0.23768958449363708
step: 60, loss: 0.3132242262363434
step: 70, loss: 0.3320189118385315
step: 80, loss: 0.07391687482595444
step: 90, loss: 0.6273192763328552
step: 100, loss: 0.2547493278980255
step: 110, loss: 0.25394460558891296
step: 120, loss: 0.12581779062747955
step: 130, loss: 0.2536751329898834
step: 140, loss: 0.17465807497501373
step: 150, loss: 0.30880340933799744
step: 160, loss: 0.24868103861808777
step: 170, loss: 0.3767927885055542
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.19509851033157133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31045106053352356
step: 10, loss: 0.11362604051828384
step: 20, loss: 0.37092843651771545
step: 30, loss: 0.24456791579723358
step: 40, loss: 0.44965818524360657
step: 50, loss: 0.2576751708984375
step: 60, loss: 0.4094502031803131

```

```

step: 70, loss: 0.3781326413154602
step: 80, loss: 0.19025103747844696
step: 90, loss: 0.3218136727809906
step: 100, loss: 0.1870298683643341
step: 110, loss: 0.4535766839981079
step: 120, loss: 0.4419199526309967
step: 130, loss: 0.4870277941226959
step: 140, loss: 0.25056231021881104
step: 150, loss: 0.16209495067596436
step: 160, loss: 0.3117968738079071
step: 170, loss: 0.23809745907783508
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.19509851033157133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24171456694602966
step: 10, loss: 0.45067647099494934
step: 20, loss: 0.45677441358566284
step: 30, loss: 0.5190474987030029
step: 40, loss: 0.12999822199344635
step: 50, loss: 0.4298291802406311
step: 60, loss: 0.5359781980514526
step: 70, loss: 0.3607417941093445
step: 80, loss: 0.1724051982164383
step: 90, loss: 0.4469091296195984
step: 100, loss: 0.3056996166706085
step: 110, loss: 0.3683735430240631
step: 120, loss: 0.3569650948047638
step: 130, loss: 0.16821564733982086
step: 140, loss: 0.09507281333208084
step: 150, loss: 0.3203033208847046
step: 160, loss: 0.37840932607650757
step: 170, loss: 0.31282490491867065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24764468371467024, f1=0.22811671087533159,
best_f1=0.22811671087533159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4394848644733429
step: 10, loss: 0.5089384913444519
step: 20, loss: 0.24146701395511627
step: 30, loss: 0.30201414227485657
step: 40, loss: 0.2322951704263687
step: 50, loss: 0.36524486541748047
step: 60, loss: 0.35631486773490906
step: 70, loss: 0.229348286986351
step: 80, loss: 0.3112145662307739
step: 90, loss: 0.49558889865875244
step: 100, loss: 0.23744793236255646
step: 110, loss: 0.5244325399398804
step: 120, loss: 0.3239734470844269
step: 130, loss: 0.3178585469722748
step: 140, loss: 0.4608125388622284
step: 150, loss: 0.2722031772136688
step: 160, loss: 0.15686921775341034
step: 170, loss: 0.45353463292121887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.284688995215311, f1=0.25806451612903225,
best_f1=0.25806451612903225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2872990071773529
step: 10, loss: 0.5502951145172119
step: 20, loss: 0.37323930859565735
step: 30, loss: 0.14938588440418243
step: 40, loss: 0.3815414309501648
step: 50, loss: 0.20372554659843445
step: 60, loss: 0.2997150421142578
step: 70, loss: 0.3606926500797272
step: 80, loss: 0.15151533484458923
step: 90, loss: 0.4198896884918213
step: 100, loss: 0.47646480798721313
step: 110, loss: 0.3426738679409027
step: 120, loss: 0.41460666060447693
step: 130, loss: 0.28031203150749207
step: 140, loss: 0.34792497754096985
step: 150, loss: 0.3438016474246979
step: 160, loss: 0.25117018818855286

```

```

step: 170, loss: 0.4078541398048401
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.28835489833641403, f1=0.23970037453183518,
best_f1=0.23970037453183518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4369175434112549
step: 10, loss: 0.2812469005584717
step: 20, loss: 0.23564413189888
step: 30, loss: 0.6073693037033081
step: 40, loss: 0.34798431396484375
step: 50, loss: 0.34104984998703003
step: 60, loss: 0.42365768551826477
step: 70, loss: 0.2539379894733429
step: 80, loss: 0.27211496233940125
step: 90, loss: 0.2252773493528366
step: 100, loss: 0.21758247911930084
step: 110, loss: 0.24068138003349304
step: 120, loss: 0.26480832695961
step: 130, loss: 0.12691615521907806
step: 140, loss: 0.3144363462924957
step: 150, loss: 0.3029501438140869
step: 160, loss: 0.22388306260108948
step: 170, loss: 0.39685696363449097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3054331864904552, f1=0.25852272727272724,
best_f1=0.25852272727272724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3934549391269684
step: 10, loss: 0.2658974528312683
step: 20, loss: 0.17762893438339233
step: 30, loss: 0.44905853271484375
step: 40, loss: 0.20726509392261505

```

```

step: 50, loss: 0.1519460380077362
step: 60, loss: 0.35223308205604553
step: 70, loss: 0.26750531792640686
step: 80, loss: 0.18342231214046478
step: 90, loss: 0.34988051652908325
step: 100, loss: 0.4269312918186188
step: 110, loss: 0.3181116580963135
step: 120, loss: 0.40747809410095215
step: 130, loss: 0.3897997736930847
step: 140, loss: 0.4917646646499634
step: 150, loss: 0.3763106167316437
step: 160, loss: 0.23177723586559296
step: 170, loss: 0.14548908174037933
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45177664974619286, f1=0.418848167539267,
best_f1=0.418848167539267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34818580746650696
step: 10, loss: 0.252633661031723
step: 20, loss: 0.15402118861675262
step: 30, loss: 0.2730884253978729
step: 40, loss: 0.21346242725849152
step: 50, loss: 0.13716651499271393
step: 60, loss: 0.2541043758392334
step: 70, loss: 0.18314093351364136
step: 80, loss: 0.1781444549560547
step: 90, loss: 0.16486920416355133
step: 100, loss: 0.1605304628610611
step: 110, loss: 0.2629065215587616
step: 120, loss: 0.16050441563129425
step: 130, loss: 0.2518408000469208
step: 140, loss: 0.13923193514347076
step: 150, loss: 0.10592562705278397
step: 160, loss: 0.34765979647636414
step: 170, loss: 0.36396968364715576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7481662591687042, f1=0.7848699763593381,
best_f1=0.7848699763593381
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15256454050540924
step: 10, loss: 0.0437600277364254
step: 20, loss: 0.12264621257781982
step: 30, loss: 0.22168055176734924
step: 40, loss: 0.24100081622600555
step: 50, loss: 0.14644353091716766
step: 60, loss: 0.11409671604633331
step: 70, loss: 0.4024983048439026
step: 80, loss: 0.056299418210983276
step: 90, loss: 0.21942003071308136
step: 100, loss: 0.07679345458745956
step: 110, loss: 0.06079447269439697
step: 120, loss: 0.03590155020356178
step: 130, loss: 0.09686533361673355
step: 140, loss: 0.17862039804458618
step: 150, loss: 0.1707839071750641
step: 160, loss: 0.2458382546901703
step: 170, loss: 0.036441683769226074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7736842105263158, f1=0.8109452736318408,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13591726124286652
step: 10, loss: 0.0622139535844326
step: 20, loss: 0.14409968256950378
step: 30, loss: 0.14361073076725006
step: 40, loss: 0.09267888218164444
step: 50, loss: 0.020652739331126213
step: 60, loss: 0.12286754697561264
step: 70, loss: 0.058382607996463776
step: 80, loss: 0.029188813641667366
step: 90, loss: 0.0358792208135128
step: 100, loss: 0.1942116767168045

```

```

step: 110, loss: 0.12081395834684372
step: 120, loss: 0.07683485746383667
step: 130, loss: 0.1364748477935791
step: 140, loss: 0.06258697062730789
step: 150, loss: 0.07330258190631866
step: 160, loss: 0.20412221550941467
step: 170, loss: 0.1323481947183609
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7840375586854459, f1=0.7877461706783371,
best_f1=0.7877461706783371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026065926998853683
step: 10, loss: 0.07821966707706451
step: 20, loss: 0.024715131148695946
step: 30, loss: 0.08543524146080017
step: 40, loss: 0.004769920837134123
step: 50, loss: 0.07510899752378464
step: 60, loss: 0.010077297687530518
step: 70, loss: 0.22654062509536743
step: 80, loss: 0.039461567997932434
step: 90, loss: 0.14541180431842804
step: 100, loss: 0.2209073305130005
step: 110, loss: 0.07369574904441833
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.026296531781554222
step: 130, loss: 0.1896834373474121
step: 140, loss: 0.07882322371006012
step: 150, loss: 0.012008250690996647
step: 160, loss: 0.05901148542761803
step: 170, loss: 0.17800599336624146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.7898734177215191, f1=0.827250608272506,
best_f1=0.827250608272506
mv: cannot stat '*_dev.pt': No such file or directory

```


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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
1916it [00:06, 310.13it/s]
load_f1 = 0.5346938775510204
real_f1 = 0.5175600739371534
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, 187.99it/s]
```

4.3 DITTO DIRTY

4.3.1 DBLP-ACM - Running the matcher

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

```

--task Dirty/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.5638386011123657
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4523698687553406
step: 20, loss: 0.5943716764450073
step: 30, loss: 0.32205304503440857
step: 40, loss: 0.2811664044857025
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.3941330015659332
step: 60, loss: 0.15165716409683228
step: 70, loss: 0.149578258395195
step: 80, loss: 0.0649687871336937
step: 90, loss: 0.09953347593545914
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 100, loss: 0.28895601630210876
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 110, loss: 0.09085365384817123
step: 120, loss: 0.079067662358284
step: 130, loss: 0.15373997390270233
step: 140, loss: 0.019916338846087456
step: 150, loss: 0.13686060905456543
step: 160, loss: 0.02022949792444706
step: 170, loss: 0.10317860543727875
step: 180, loss: 0.1300976425409317
step: 190, loss: 0.035699598491191864
step: 200, loss: 0.025361714884638786
step: 210, loss: 0.025572264567017555
step: 220, loss: 0.12259423732757568
step: 230, loss: 0.005914803594350815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9642857142857144, f1=0.971687429218573,
best_f1=0.971687429218573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0039229257963597775
step: 10, loss: 0.09988950937986374
step: 20, loss: 0.014446964487433434
step: 30, loss: 0.011932021006941795
step: 40, loss: 0.019484175369143486
step: 50, loss: 0.009907406754791737
step: 60, loss: 0.004133705515414476
step: 70, loss: 0.0241899024695158
step: 80, loss: 0.017749611288309097
step: 90, loss: 0.04406895861029625
step: 100, loss: 0.0009312891634181142
step: 110, loss: 0.020699385553598404
step: 120, loss: 0.012304089963436127
step: 130, loss: 0.018607843667268753
step: 140, loss: 0.009833796881139278
step: 150, loss: 0.0781368538737297
step: 160, loss: 0.022741515189409256
step: 170, loss: 0.0011244738707318902
step: 180, loss: 0.00954620074480772
step: 190, loss: 0.007966360077261925
step: 200, loss: 0.020138736814260483
step: 210, loss: 0.05118654668331146
step: 220, loss: 0.0043209632858633995
step: 230, loss: 0.006040541455149651
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753914988814317, f1=0.9727891156462585,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028138870373368263
step: 10, loss: 0.0064546167850494385

```

```

step: 20, loss: 0.02040894888341427
step: 30, loss: 0.007111228536814451
step: 40, loss: 0.06290104240179062
step: 50, loss: 0.04505313187837601
step: 60, loss: 0.020750729367136955
step: 70, loss: 0.0072885374538600445
step: 80, loss: 0.04730507731437683
step: 90, loss: 0.0063994573429226875
step: 100, loss: 0.0037176255136728287
step: 110, loss: 0.04694876819849014
step: 120, loss: 0.0010777291608974338
step: 130, loss: 0.051488690078258514
step: 140, loss: 0.004397118929773569
step: 150, loss: 0.05093053728342056
step: 160, loss: 0.0013061942299827933
step: 170, loss: 0.0009630421991460025
step: 180, loss: 0.0036893479991704226
step: 190, loss: 0.004109215456992388
step: 200, loss: 0.00524417869746685
step: 210, loss: 0.00297450739890337
step: 220, loss: 0.18423093855381012
step: 230, loss: 0.018982985988259315
/usr/local/lib/python3.7/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=0.9688195991091313, f1=0.968609865470852,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028713420033454895
step: 10, loss: 0.008731655776500702
step: 20, loss: 0.010256528854370117
step: 30, loss: 0.00821844395250082
step: 40, loss: 0.15295520424842834
step: 50, loss: 0.02324007824063301
step: 60, loss: 0.04432080313563347
step: 70, loss: 0.009164699353277683
step: 80, loss: 0.10067487508058548
step: 90, loss: 0.05881055071949959
step: 100, loss: 0.030152132734656334
step: 110, loss: 0.0008258325979113579
step: 120, loss: 0.004054793156683445
step: 130, loss: 0.008922174572944641
step: 140, loss: 0.0061120400205254555
step: 150, loss: 0.0003503590414766222
step: 160, loss: 0.019353248178958893
step: 170, loss: 0.014148274436593056

```

```

step: 180, loss: 0.005203989800065756
step: 190, loss: 0.001507187495008111
step: 200, loss: 0.15500666201114655
step: 210, loss: 0.004026660695672035
step: 220, loss: 0.005589818581938744
step: 230, loss: 0.022959917783737183
/usr/local/lib/python3.7/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=0.9741282339707535, f1=0.9607182940516273,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011314123403280973
step: 10, loss: 0.0028423615731298923
step: 20, loss: 0.12356982380151749
step: 30, loss: 0.010613247752189636
step: 40, loss: 0.0022879713214933872
step: 50, loss: 0.0008944576256908476
step: 60, loss: 0.1055261492729187
step: 70, loss: 0.004661455750465393
step: 80, loss: 0.011240310035645962
step: 90, loss: 0.03648453205823898
step: 100, loss: 0.004764285869896412
step: 110, loss: 0.0007651667692698538
step: 120, loss: 0.0011033033952116966
step: 130, loss: 0.0004997379146516323
step: 140, loss: 0.0017065175343304873
step: 150, loss: 0.018099121749401093
step: 160, loss: 0.0009027851046994328
step: 170, loss: 0.0004492372099775821
step: 180, loss: 0.009134571067988873
step: 190, loss: 0.02304721251130104
step: 200, loss: 0.0010214895009994507
step: 210, loss: 0.05766266584396362
step: 220, loss: 0.004509145859628916
step: 230, loss: 0.09639514982700348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.9627959413754228,
best_f1=0.9627959413754228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009530344977974892
step: 10, loss: 0.023411085829138756
step: 20, loss: 0.0018176346784457564
step: 30, loss: 0.0018784867133945227
step: 40, loss: 0.0001393311977153644
step: 50, loss: 0.0008947452297434211
step: 60, loss: 0.009506314061582088
step: 70, loss: 0.10812008380889893
step: 80, loss: 0.0059811570681631565
step: 90, loss: 0.003744247369468212
step: 100, loss: 0.0022882830817252398
step: 110, loss: 0.017013438045978546
step: 120, loss: 0.00033668920514173806
step: 130, loss: 0.0030260630883276463
step: 140, loss: 0.00432936754077673
step: 150, loss: 0.0006080649909563363
step: 160, loss: 0.02230977825820446
step: 170, loss: 0.0015792545164003968
step: 180, loss: 0.0020638068672269583
step: 190, loss: 0.0020487322472035885
step: 200, loss: 0.0014485330320894718
step: 210, loss: 0.005101310089230537
step: 220, loss: 0.009511835873126984
step: 230, loss: 0.0001861050695879385
/usr/local/lib/python3.7/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=0.9797297297297298, f1=0.967452300785634,
best_f1=0.9627959413754228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005539256962947547
step: 10, loss: 0.0417775921523571
step: 20, loss: 0.00045473669888451695
step: 30, loss: 9.707578283268958e-05
step: 40, loss: 0.0002876034704968333
step: 50, loss: 0.0001886599784484133
step: 60, loss: 0.00021992460824549198
step: 70, loss: 0.00015043398889247328
step: 80, loss: 0.00014683231711387634
step: 90, loss: 0.00484359310939908
step: 100, loss: 0.0003707570140250027
step: 110, loss: 0.00016181854880414903
step: 120, loss: 0.0032959552481770515
step: 130, loss: 0.11230093240737915

```

```

step: 140, loss: 0.0001886206737253815
step: 150, loss: 0.11338454484939575
step: 160, loss: 0.001316013396717608
step: 170, loss: 0.016936402767896652
step: 180, loss: 0.0014085503062233329
step: 190, loss: 0.02640402689576149
step: 200, loss: 0.0009014446404762566
step: 210, loss: 0.07797650247812271
step: 220, loss: 0.004454297479242086
step: 230, loss: 0.005560419522225857
/usr/local/lib/python3.7/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=0.978675645342312, f1=0.961625282167043,
best_f1=0.9627959413754228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027478947304189205
step: 10, loss: 0.025405049324035645
step: 20, loss: 0.0022652577608823776
step: 30, loss: 0.0006364356959238648
step: 40, loss: 0.008178996853530407
step: 50, loss: 0.0009432208025828004
step: 60, loss: 0.0007345497142523527
step: 70, loss: 0.00030832039192318916
step: 80, loss: 0.0029381413478404284
step: 90, loss: 0.00043679174268618226
step: 100, loss: 0.0010112974559888244
step: 110, loss: 0.013435759581625462
step: 120, loss: 0.0026455316692590714
step: 130, loss: 0.005952590610831976
step: 140, loss: 0.013653111644089222
step: 150, loss: 0.14773088693618774
step: 160, loss: 0.001068608369678259
step: 170, loss: 0.05101301521062851
step: 180, loss: 8.492010965710506e-05
step: 190, loss: 0.0016383351758122444
step: 200, loss: 0.00809935387223959
step: 210, loss: 0.017304331064224243
step: 220, loss: 0.0010782623430714011
step: 230, loss: 0.0005355398752726614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.980963045912654, f1=0.9730941704035874,
best_f1=0.9730941704035874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018853724468499422
step: 10, loss: 0.00039622950134798884
step: 20, loss: 0.0005628018407151103
step: 30, loss: 0.00017807212134357542
step: 40, loss: 0.00019596857600845397
step: 50, loss: 0.00016670936020091176
step: 60, loss: 0.0004030083364341408
step: 70, loss: 0.008992655202746391
step: 80, loss: 8.434526534983888e-05
step: 90, loss: 0.045768603682518005
step: 100, loss: 0.0006603896617889404
step: 110, loss: 0.00036098677082918584
step: 120, loss: 0.037099916487932205
step: 130, loss: 0.0021421508863568306
step: 140, loss: 0.0006951014511287212
step: 150, loss: 0.0002476138179190457
step: 160, loss: 0.03692522644996643
step: 170, loss: 0.00011076157534262165
step: 180, loss: 0.007242790423333645
step: 190, loss: 5.947998579358682e-05
step: 200, loss: 0.0006915254634805024
step: 210, loss: 0.030305184423923492
step: 220, loss: 5.817749843117781e-05
step: 230, loss: 0.00537197245284915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832402234636871, f1=0.9711111111111111,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006491499952971935
step: 10, loss: 0.00025756063405424356
step: 20, loss: 0.0522838793694973
step: 30, loss: 0.0008537247776985168
step: 40, loss: 0.00041944000986404717
step: 50, loss: 0.0006461867596954107

```

```

step: 60, loss: 0.0002844557457137853
step: 70, loss: 0.008806722238659859
step: 80, loss: 0.0007206479203887284
step: 90, loss: 0.0005016544018872082
step: 100, loss: 0.0003554314316716045
step: 110, loss: 0.006514484528452158
step: 120, loss: 9.52857153606601e-05
step: 130, loss: 6.782682612538338e-05
step: 140, loss: 0.0015306829009205103
step: 150, loss: 6.709124863846228e-05
step: 160, loss: 3.2445612305309623e-05
step: 170, loss: 4.733840614790097e-05
step: 180, loss: 0.00028384773759171367
step: 190, loss: 5.6981090892804787e-05
step: 200, loss: 0.00010336384002584964
step: 210, loss: 0.0003417646512389183
step: 220, loss: 0.010023909620940685
step: 230, loss: 6.28947964287363e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9832402234636871, f1=0.9740698985343857,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.249508780660108e-05
step: 10, loss: 0.00010904873488470912
step: 20, loss: 0.00898087490350008
step: 30, loss: 0.00022130351862870157
step: 40, loss: 2.76892078545643e-05
step: 50, loss: 0.005271488334983587
step: 60, loss: 4.195979636278935e-05
step: 70, loss: 9.791512275114655e-05
step: 80, loss: 0.008579403162002563
step: 90, loss: 0.06750571727752686
step: 100, loss: 0.00012364033318590373
step: 110, loss: 0.0033069003839045763
step: 120, loss: 6.190973363118246e-05
step: 130, loss: 6.756982475053519e-05
step: 140, loss: 0.0004895308520644903
step: 150, loss: 0.00013706745812669396
step: 160, loss: 0.049632396548986435
step: 170, loss: 0.00032001116778701544
step: 180, loss: 4.3803378503071144e-05
step: 190, loss: 6.034086618456058e-05
step: 200, loss: 0.0007086555124260485
step: 210, loss: 3.945243588532321e-05

```

```

step: 220, loss: 0.00014863081742078066
step: 230, loss: 0.00045997879351489246
/usr/local/lib/python3.7/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=0.972972972972973,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004106390289962292
step: 10, loss: 3.035246118088253e-05
step: 20, loss: 0.008016817271709442
step: 30, loss: 0.02484768256545067
step: 40, loss: 4.683802035287954e-05
step: 50, loss: 0.000713006651494652
step: 60, loss: 0.002023642882704735
step: 70, loss: 8.290139521704987e-05
step: 80, loss: 2.3047727154335007e-05
step: 90, loss: 3.48784560628701e-05
step: 100, loss: 2.6366747988504358e-05
step: 110, loss: 3.774517972487956e-05
step: 120, loss: 0.0009120141621679068
step: 130, loss: 0.00026200167485512793
step: 140, loss: 0.00091278221225366
step: 150, loss: 5.008328298572451e-05
step: 160, loss: 0.008341684006154537
step: 170, loss: 0.02050573006272316
step: 180, loss: 0.0006398113328032196
step: 190, loss: 0.0003219414793420583
step: 200, loss: 5.785351822851226e-05
step: 210, loss: 0.03307061269879341
step: 220, loss: 0.009577676653862
step: 230, loss: 7.754421676509082e-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=0.9842342342342343, f1=0.9688195991091313,
best_f1=0.9688195991091313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024278820492327213
step: 10, loss: 0.005436259787529707

```

```

step: 20, loss: 3.597027171053924e-05
step: 30, loss: 2.620275699882768e-05
step: 40, loss: 4.684666055254638e-05
step: 50, loss: 0.002899142215028405
step: 60, loss: 3.0207671443349682e-05
step: 70, loss: 0.00043037926661781967
step: 80, loss: 0.00010376393038313836
step: 90, loss: 2.873599987651687e-05
step: 100, loss: 0.00025326863396912813
step: 110, loss: 6.925783236511052e-05
step: 120, loss: 0.00011301122140139341
step: 130, loss: 7.767554052406922e-05
step: 140, loss: 0.0013271105708554387
step: 150, loss: 0.0010204671416431665
step: 160, loss: 0.000518325250595808
step: 170, loss: 3.368299439898692e-05
step: 180, loss: 0.012961437925696373
step: 190, loss: 3.916263085557148e-05
step: 200, loss: 0.00016182553372345865
step: 210, loss: 9.863745071925223e-05
step: 220, loss: 4.089974027010612e-05
step: 230, loss: 5.872430119779892e-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=0.9853768278965129, f1=0.972972972972973,
best_f1=0.972972972972973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.230074308812618e-05
step: 10, loss: 7.240369450300932e-05
step: 20, loss: 0.0022431141696870327
step: 30, loss: 3.948892481275834e-05
step: 40, loss: 3.410716453799978e-05
step: 50, loss: 3.87026448152028e-05
step: 60, loss: 2.599440995254554e-05
step: 70, loss: 5.015286660636775e-05
step: 80, loss: 0.0003334767825435847
step: 90, loss: 0.0001334834232693538
step: 100, loss: 0.00028011976974084973
step: 110, loss: 0.00029639541753567755
step: 120, loss: 1.80822353286203e-05
step: 130, loss: 0.00018017903494182974

```

```

step: 140, loss: 4.316839840612374e-05
step: 150, loss: 1.719940519251395e-05
step: 160, loss: 3.152465433231555e-05
step: 170, loss: 0.00015280974912457168
step: 180, loss: 2.781596231216099e-05
step: 190, loss: 0.00010085992107633501
step: 200, loss: 2.5782117518247105e-05
step: 210, loss: 4.8727462854003534e-05
step: 220, loss: 0.0001125984126701951
step: 230, loss: 2.3993798095034435e-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=0.9831649831649831, f1=0.9741282339707535,
best_f1=0.972972972972973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00273455074056983
step: 10, loss: 2.7905443857889622e-05
step: 20, loss: 5.230960960034281e-05
step: 30, loss: 3.3030719350790605e-05
step: 40, loss: 0.0001433771540177986
step: 50, loss: 2.2418296794057824e-05
step: 60, loss: 0.008326292037963867
step: 70, loss: 0.00011185206676600501
step: 80, loss: 5.285022780299187e-05
step: 90, loss: 2.6031440938822925e-05
step: 100, loss: 2.1684512830688618e-05
step: 110, loss: 0.00015729412552900612
step: 120, loss: 0.000380361860152334
step: 130, loss: 2.5279186957050115e-05
step: 140, loss: 0.028418561443686485
step: 150, loss: 0.016904566437005997
step: 160, loss: 0.00017276551807299256
step: 170, loss: 1.5716814232291654e-05
step: 180, loss: 2.757391303020995e-05
step: 190, loss: 0.00861378014087677
step: 200, loss: 0.0008563771261833608
step: 210, loss: 0.00010452191054355353
step: 220, loss: 0.04774617403745651
step: 230, loss: 2.6922029064735398e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.984304932735426, f1=0.9709821428571428,
best_f1=0.972972972972973
mv: cannot stat '*_dev.pt': No such file or directory

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

2473it [00:12, 199.34it/s]

load_f1 = 0.9831649831649831

real_f1 = 0.9809203142536477

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

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

4398it [00:24, 182.43it/s]

4.3.2 DBLP-GoogleScholar - Running the matcher

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

```

--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

```

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

Step 3: run the matcher

```

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

```

Downloading: 100% 481/481 [00:00<00:00, 652kB/s]
 Downloading: 100% 899k/899k [00:00<00:00, 23.4MB/s]
 Downloading: 100% 456k/456k [00:00<00:00, 21.1MB/s]
 Downloading: 100% 501M/501M [00:06<00:00, 73.1MB/s]
 Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16

```

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.7174947261810303
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4449824094772339
step: 20, loss: 0.2804468870162964
step: 30, loss: 0.33779892325401306
step: 40, loss: 0.39265310764312744
step: 50, loss: 0.7488257884979248
step: 60, loss: 0.37830686569213867
step: 70, loss: 0.46227577328681946
step: 80, loss: 0.49813312292099
step: 90, loss: 0.33703479170799255
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.3160206079483032
step: 110, loss: 0.11733493953943253
step: 120, loss: 0.29762789607048035
step: 130, loss: 0.11998412013053894
step: 140, loss: 0.2610205411911011
step: 150, loss: 0.02174721658229828
step: 160, loss: 0.08615370839834213
step: 170, loss: 0.11679836362600327
step: 180, loss: 0.14203999936580658
step: 190, loss: 0.14383889734745026
step: 200, loss: 0.041643209755420685
step: 210, loss: 0.08021432161331177
step: 220, loss: 0.03884726017713547
step: 230, loss: 0.26130762696266174
step: 240, loss: 0.04301196709275246
step: 250, loss: 0.08685964345932007

```



```

step: 260, loss: 0.15705157816410065
step: 270, loss: 0.3417229950428009
step: 280, loss: 0.088765949010849
step: 290, loss: 0.15115827322006226
step: 300, loss: 0.05309642478823662
step: 310, loss: 0.08456456661224365
step: 320, loss: 0.09385097026824951
step: 330, loss: 0.09659477323293686
step: 340, loss: 0.4262617826461792
step: 350, loss: 0.0829811617732048
step: 360, loss: 0.05526850372552872
step: 370, loss: 0.13068179786205292
step: 380, loss: 0.21021413803100586
step: 390, loss: 0.04910207539796829
step: 400, loss: 0.05459132790565491
step: 410, loss: 0.26743385195732117
step: 420, loss: 0.02785203419625759
step: 430, loss: 0.02743733860552311
step: 440, loss: 0.07127755880355835
step: 450, loss: 0.12427426129579544
step: 460, loss: 0.05253364518284798
step: 470, loss: 0.04015463590621948
step: 480, loss: 0.09609539061784744
step: 490, loss: 0.1974400281906128
step: 500, loss: 0.07620927691459656
step: 510, loss: 0.033242084085941315
step: 520, loss: 0.15634603798389435
step: 530, loss: 0.027915067970752716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9463955637707948, f1=0.937847866419295,
best_f1=0.937847866419295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06662846356630325
step: 10, loss: 0.07039418071508408
step: 20, loss: 0.06630813330411911
step: 30, loss: 0.04361686483025551
step: 40, loss: 0.17032495141029358
step: 50, loss: 0.04512980207800865
step: 60, loss: 0.09459268301725388
step: 70, loss: 0.0332869291305542

```

step: 80, loss: 0.01953289285302162
step: 90, loss: 0.002329480368643999
step: 100, loss: 0.1599183827638626
step: 110, loss: 0.017012789845466614
step: 120, loss: 0.028519073501229286
step: 130, loss: 0.004071661736816168
step: 140, loss: 0.05506123602390289
step: 150, loss: 0.029061252251267433
step: 160, loss: 0.039982669055461884
step: 170, loss: 0.04918700456619263
step: 180, loss: 0.018934864550828934
step: 190, loss: 0.03591708466410637
step: 200, loss: 0.3211357891559601
step: 210, loss: 0.050431061536073685
step: 220, loss: 0.0014417769853025675
step: 230, loss: 0.1962776929140091
step: 240, loss: 0.03674856945872307
step: 250, loss: 0.03748684003949165
step: 260, loss: 0.10181724280118942
step: 270, loss: 0.06666616350412369
step: 280, loss: 0.13120439648628235
step: 290, loss: 0.03103785403072834
step: 300, loss: 0.035520218312740326
step: 310, loss: 0.030648738145828247
step: 320, loss: 0.09407962113618851
step: 330, loss: 0.03157272934913635
step: 340, loss: 0.040109895169734955
step: 350, loss: 0.005899844691157341
step: 360, loss: 0.1340167075395584
step: 370, loss: 0.027176126837730408
step: 380, loss: 0.11970342695713043
step: 390, loss: 0.007393470499664545
step: 400, loss: 0.14441163837909698
step: 410, loss: 0.039082154631614685
step: 420, loss: 0.017247891053557396
step: 430, loss: 0.18102186918258667
step: 440, loss: 0.005844161845743656
step: 450, loss: 0.05956905707716942
step: 460, loss: 0.10849976539611816
step: 470, loss: 0.07201803475618362
step: 480, loss: 0.026008404791355133
step: 490, loss: 0.05573197081685066
step: 500, loss: 0.0032697892747819424
step: 510, loss: 0.028895987197756767
step: 520, loss: 0.285699725151062
step: 530, loss: 0.050535861402750015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9530386740331491, f1=0.9466911764705882,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12813852727413177
step: 10, loss: 0.046609409153461456
step: 20, loss: 0.03346817567944527
step: 30, loss: 0.1259748637676239
step: 40, loss: 0.07825610041618347
step: 50, loss: 0.012851813808083534
step: 60, loss: 0.018292104825377464
step: 70, loss: 0.01583663374185562
step: 80, loss: 0.16596300899982452
step: 90, loss: 0.03825230151414871
step: 100, loss: 0.022170396521687508
step: 110, loss: 0.0026635031681507826
step: 120, loss: 0.2127196341753006
step: 130, loss: 0.04547547921538353
step: 140, loss: 0.022261686623096466
step: 150, loss: 0.016338098794221878
step: 160, loss: 0.01996716856956482
step: 170, loss: 0.005240068770945072
step: 180, loss: 0.04181993007659912
step: 190, loss: 0.010163750499486923
step: 200, loss: 0.018644705414772034
step: 210, loss: 0.02632778324186802
step: 220, loss: 0.147396519780159
step: 230, loss: 0.07786824554204941
step: 240, loss: 0.2338661104440689
step: 250, loss: 0.03526771068572998
step: 260, loss: 0.09810104966163635
step: 270, loss: 0.08926811069250107
step: 280, loss: 0.0014028125442564487
step: 290, loss: 0.006189560983330011
step: 300, loss: 0.16725625097751617
step: 310, loss: 0.09161274135112762
step: 320, loss: 0.029308242723345757
step: 330, loss: 0.007692490238696337
step: 340, loss: 0.004999042022973299
step: 350, loss: 0.13737209141254425
step: 360, loss: 0.00897139124572277
step: 370, loss: 0.04205271601676941

```

```
step: 380, loss: 0.05979572981595993
step: 390, loss: 0.01792207546532154
step: 400, loss: 0.16703073680400848
step: 410, loss: 0.009511658921837807
step: 420, loss: 0.01033011544495821
step: 430, loss: 0.056664079427719116
step: 440, loss: 0.21319200098514557
step: 450, loss: 0.08325961977243423
step: 460, loss: 0.07365136593580246
step: 470, loss: 0.019779277965426445
step: 480, loss: 0.022973740473389626
step: 490, loss: 0.005354771856218576
step: 500, loss: 0.042092420160770416
step: 510, loss: 0.08498226851224899
step: 520, loss: 0.006399867124855518
step: 530, loss: 0.01848386786878109
/usr/local/lib/python3.7/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=0.9453526389537599, f1=0.943378568086102,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04011860489845276
step: 10, loss: 0.002041485393419862
step: 20, loss: 0.05360068008303642
step: 30, loss: 0.11259707808494568
step: 40, loss: 0.015807606279850006
step: 50, loss: 0.04508981481194496
step: 60, loss: 0.015544977970421314
step: 70, loss: 0.0839652568101883
step: 80, loss: 0.08991718292236328
step: 90, loss: 0.07088813185691833
step: 100, loss: 0.0016314475797116756
step: 110, loss: 0.15775562822818756
step: 120, loss: 0.009909240528941154
step: 130, loss: 0.04289039969444275
step: 140, loss: 0.1992596536874771
step: 150, loss: 0.004642983432859182
step: 160, loss: 0.0061562261544167995
step: 170, loss: 0.08651769161224365
step: 180, loss: 0.12309186160564423
step: 190, loss: 0.034514833241701126
step: 200, loss: 0.07174952328205109
step: 210, loss: 0.0012983577325940132
step: 220, loss: 0.02181163989007473
step: 230, loss: 0.006348251830786467
```

```

step: 240, loss: 0.016908785328269005
step: 250, loss: 0.02271391823887825
step: 260, loss: 0.0010114518227055669
step: 270, loss: 0.026851089671254158
step: 280, loss: 0.0067167095839977264
step: 290, loss: 0.03223968669772148
step: 300, loss: 0.030136262997984886
step: 310, loss: 0.0034949134569615126
step: 320, loss: 0.11204579472541809
step: 330, loss: 0.004900939296931028
step: 340, loss: 0.006868863478302956
step: 350, loss: 0.21606290340423584
step: 360, loss: 0.06334969401359558
step: 370, loss: 0.0018583323108032346
step: 380, loss: 0.012407374568283558
step: 390, loss: 0.0017460761591792107
step: 400, loss: 0.002987045096233487
step: 410, loss: 0.0008664357010275126
step: 420, loss: 0.010006014257669449
step: 430, loss: 0.010375839658081532
step: 440, loss: 0.005892714951187372
step: 450, loss: 0.008212458342313766
step: 460, loss: 0.007980819791555405
step: 470, loss: 0.003163154935464263
step: 480, loss: 0.01946609653532505
step: 490, loss: 0.03913724049925804
step: 500, loss: 0.13862906396389008
step: 510, loss: 0.06682979315519333
step: 520, loss: 0.04484154284000397
step: 530, loss: 0.1344504952430725
/usr/local/lib/python3.7/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=0.9489461358313818, f1=0.9499298081422555,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007190250908024609
step: 10, loss: 0.018922820687294006
step: 20, loss: 0.0020006978884339333
step: 30, loss: 0.1526477187871933
step: 40, loss: 0.037585075944662094
step: 50, loss: 0.11880020052194595
step: 60, loss: 0.10828684270381927
step: 70, loss: 0.054724179208278656
step: 80, loss: 0.034895338118076324
step: 90, loss: 0.22583788633346558

```

```
step: 100, loss: 0.07996881753206253
step: 110, loss: 0.009307480417191982
step: 120, loss: 0.04238782823085785
step: 130, loss: 0.0012286584824323654
step: 140, loss: 0.011189102195203304
step: 150, loss: 0.016593551263213158
step: 160, loss: 0.008548008278012276
step: 170, loss: 0.2084297090768814
step: 180, loss: 0.03172443434596062
step: 190, loss: 0.018631847575306892
step: 200, loss: 0.0039008022285997868
step: 210, loss: 0.0009866405744105577
step: 220, loss: 0.01254820916801691
step: 230, loss: 0.005275403149425983
step: 240, loss: 0.03856964409351349
step: 250, loss: 0.04817376285791397
step: 260, loss: 0.0005386040429584682
step: 270, loss: 0.0059822434559464455
step: 280, loss: 0.008454675786197186
step: 290, loss: 0.01816342957317829
step: 300, loss: 0.058318160474300385
step: 310, loss: 0.017739461734890938
step: 320, loss: 0.10525574535131454
step: 330, loss: 0.033260613679885864
step: 340, loss: 0.055447060614824295
step: 350, loss: 0.0020815986208617687
step: 360, loss: 0.0003278263611719012
step: 370, loss: 0.00035716165439225733
step: 380, loss: 0.001334059634245932
step: 390, loss: 0.005252726376056671
step: 400, loss: 0.061683688312768936
step: 410, loss: 0.035892609506845474
step: 420, loss: 0.22942900657653809
step: 430, loss: 0.005958253052085638
step: 440, loss: 0.012170503847301006
step: 450, loss: 0.010884671472012997
step: 460, loss: 0.0062895407900214195
step: 470, loss: 0.00525313476100564
step: 480, loss: 0.17536644637584686
step: 490, loss: 0.057328108698129654
step: 500, loss: 0.021828768774867058
step: 510, loss: 0.0030565145425498486
step: 520, loss: 0.038099199533462524
step: 530, loss: 0.03965415805578232
/usr/local/lib/python3.7/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=0.9469732519943688, f1=0.9407894736842105,
```

```
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05330629274249077
step: 10, loss: 0.030378155410289764
step: 20, loss: 0.041273217648267746
step: 30, loss: 0.00042645656503736973
step: 40, loss: 0.0005027953302487731
step: 50, loss: 0.00032599776750430465
step: 60, loss: 0.00033295282628387213
step: 70, loss: 0.005844191648066044
step: 80, loss: 0.0004958448698744178
step: 90, loss: 0.004777100868523121
step: 100, loss: 0.004113789647817612
step: 110, loss: 0.003060555551201105
step: 120, loss: 0.016396306455135345
step: 130, loss: 0.00638915179297328
step: 140, loss: 0.0005977555993013084
step: 150, loss: 0.010017034597694874
step: 160, loss: 0.07962377369403839
step: 170, loss: 0.013527113944292068
step: 180, loss: 0.004077235236763954
step: 190, loss: 0.06669740378856659
step: 200, loss: 0.03799892216920853
step: 210, loss: 0.004325710237026215
step: 220, loss: 0.01608295738697052
step: 230, loss: 0.0004092398739885539
step: 240, loss: 0.0028741161804646254
step: 250, loss: 0.005686810705810785
step: 260, loss: 0.0004855713923461735
step: 270, loss: 0.010415448807179928
step: 280, loss: 0.0006141734193079174
step: 290, loss: 0.00362754357047379
step: 300, loss: 0.008471321314573288
step: 310, loss: 0.01367497630417347
step: 320, loss: 0.00016594171756878495
step: 330, loss: 0.00010327380732633173
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 340, loss: 8.866116695571691e-05
step: 350, loss: 0.00061180250486359
step: 360, loss: 0.19712930917739868
step: 370, loss: 0.02826680988073349
step: 380, loss: 0.0015072865644469857
step: 390, loss: 0.005570352077484131
step: 400, loss: 0.009214511141180992
step: 410, loss: 0.00013269441842567176
step: 420, loss: 0.010693617165088654
```

```
step: 430, loss: 0.010272260755300522
step: 440, loss: 0.005360281560570002
step: 450, loss: 0.059187933802604675
step: 460, loss: 0.002333469223231077
step: 470, loss: 0.00829261727631092
step: 480, loss: 0.00918799638748169
step: 490, loss: 0.003358503570780158
step: 500, loss: 0.0005837127100676298
step: 510, loss: 0.03197629004716873
step: 520, loss: 0.0062810699455440044
step: 530, loss: 0.0016558916540816426
/usr/local/lib/python3.7/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=0.9507139567019807, f1=0.9402573529411765,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045150076039135456
step: 10, loss: 0.03393935412168503
step: 20, loss: 0.04406444728374481
step: 30, loss: 0.013944344595074654
step: 40, loss: 0.0033781102392822504
step: 50, loss: 0.0019875671714544296
step: 60, loss: 0.005492204800248146
step: 70, loss: 0.016843602061271667
step: 80, loss: 0.0043492745608091354
step: 90, loss: 0.0014118452090770006
step: 100, loss: 0.0019547822885215282
step: 110, loss: 0.0061266422271728516
step: 120, loss: 0.01712726429104805
step: 130, loss: 0.0009857920231297612
step: 140, loss: 0.0014878385700285435
step: 150, loss: 0.002235780470073223
step: 160, loss: 0.0009849787456914783
step: 170, loss: 0.0012705123517662287
step: 180, loss: 0.07180970907211304
step: 190, loss: 0.09813997149467468
step: 200, loss: 0.0033159207087010145
step: 210, loss: 0.020772816613316536
step: 220, loss: 0.099791519343853
step: 230, loss: 0.0028903826605528593
step: 240, loss: 0.034649286419153214
step: 250, loss: 0.005629264749586582
step: 260, loss: 0.01530811470001936
step: 270, loss: 0.06011362001299858
step: 280, loss: 0.004870598204433918
```



```

step: 290, loss: 0.06864164024591446
step: 300, loss: 0.0006425863248296082
step: 310, loss: 0.0008062561391852796
step: 320, loss: 0.06890639662742615
step: 330, loss: 0.0013836658326908946
step: 340, loss: 0.033621061593294144
step: 350, loss: 0.00034151884028688073
step: 360, loss: 0.11599880456924438
step: 370, loss: 0.004040147643536329
step: 380, loss: 0.0025514266453683376
step: 390, loss: 0.00034352776128798723
step: 400, loss: 0.010062674060463905
step: 410, loss: 0.013205216266214848
step: 420, loss: 0.016519447788596153
step: 430, loss: 0.005305269733071327
step: 440, loss: 0.005281772464513779
step: 450, loss: 0.003979567904025316
step: 460, loss: 0.0035376420710235834
step: 470, loss: 0.11435125768184662
step: 480, loss: 0.004244440235197544
step: 490, loss: 0.0060084727592766285
step: 500, loss: 0.024456029757857323
step: 510, loss: 0.0008231118554249406
step: 520, loss: 0.004938660655170679
step: 530, loss: 0.0016087170224636793
/usr/local/lib/python3.7/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=0.9485396383866481, f1=0.9393382352941178,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005445178248919547
step: 10, loss: 0.002619988052174449
step: 20, loss: 0.014470331370830536
step: 30, loss: 0.0012640110217034817
step: 40, loss: 9.818060061661527e-05
step: 50, loss: 0.002098489785566926
step: 60, loss: 0.0005942665738984942
step: 70, loss: 0.0024077098350971937
step: 80, loss: 0.0002765694516710937
step: 90, loss: 0.03614798188209534
step: 100, loss: 0.0038958291988819838
step: 110, loss: 0.0015810129698365927
step: 120, loss: 0.0009156573796644807
step: 130, loss: 0.000954392715357244
step: 140, loss: 0.0055544390343129635

```

```

step: 150, loss: 0.001793433679267764
step: 160, loss: 0.0003251214511692524
step: 170, loss: 0.010165512561798096
step: 180, loss: 0.0022287415340542793
step: 190, loss: 0.005152914673089981
step: 200, loss: 0.008344259113073349
step: 210, loss: 0.05768483504652977
step: 220, loss: 0.0004128810251131654
step: 230, loss: 0.044495925307273865
step: 240, loss: 0.019637450575828552
step: 250, loss: 0.0017042807303369045
step: 260, loss: 0.004975108429789543
step: 270, loss: 0.004361879080533981
step: 280, loss: 0.0023248582147061825
step: 290, loss: 0.001160992425866425
step: 300, loss: 0.0010251641506329179
step: 310, loss: 0.0072908117435872555
step: 320, loss: 0.001258320757187903
step: 330, loss: 0.0002674695278983563
step: 340, loss: 0.04990715533494949
step: 350, loss: 0.000789240060839802
step: 360, loss: 0.014516561292111874
step: 370, loss: 0.06472840160131454
step: 380, loss: 0.000351218186551705
step: 390, loss: 0.004386264365166426
step: 400, loss: 0.006188230589032173
step: 410, loss: 0.028171146288514137
step: 420, loss: 0.007374287117272615
step: 430, loss: 0.012724155560135841
step: 440, loss: 0.04562061280012131
step: 450, loss: 0.0004720105498563498
step: 460, loss: 0.00014055131759960204
step: 470, loss: 0.04300640523433685
step: 480, loss: 0.0009804927976801991
step: 490, loss: 0.0053429072722792625
step: 500, loss: 0.00011205917689949274
step: 510, loss: 0.004075071308761835
step: 520, loss: 0.00011444846313679591
step: 530, loss: 0.02897736057639122
/usr/local/lib/python3.7/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=0.951545916012921, f1=0.945287356321839,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.56019978527911e-05

```

step: 10, loss: 0.032777540385723114
step: 20, loss: 0.000942347280215472
step: 30, loss: 0.0349113792181015
step: 40, loss: 0.0004920365754514933
step: 50, loss: 0.00017006247071549296
step: 60, loss: 0.02223755605518818
step: 70, loss: 0.00932091660797596
step: 80, loss: 0.044637732207775116
step: 90, loss: 0.1038971096277237
step: 100, loss: 0.0018363728886470199
step: 110, loss: 0.02522057481110096
step: 120, loss: 0.008444871753454208
step: 130, loss: 0.0013934081653133035
step: 140, loss: 0.0008028101292438805
step: 150, loss: 0.007025551050901413
step: 160, loss: 0.0003541082551237196
step: 170, loss: 0.004576093517243862
step: 180, loss: 0.00024405476870015264
step: 190, loss: 3.249506698921323e-05
step: 200, loss: 4.671474380302243e-05
step: 210, loss: 0.0007867863751016557
step: 220, loss: 0.0022931343410164118
step: 230, loss: 0.001071558566763997
step: 240, loss: 0.0003177865582983941
step: 250, loss: 0.003669347381219268
step: 260, loss: 0.00010079316416522488
step: 270, loss: 0.0006053723627701402
step: 280, loss: 0.08691337704658508
step: 290, loss: 0.002073430921882391
step: 300, loss: 0.038025543093681335
step: 310, loss: 0.14822065830230713
step: 320, loss: 0.0016014097491279244
step: 330, loss: 0.0004766864294651896
step: 340, loss: 0.06560280174016953
step: 350, loss: 0.01692194491624832
step: 360, loss: 0.0019013938726857305
step: 370, loss: 0.012039422057569027
step: 380, loss: 0.009115018881857395
step: 390, loss: 0.000710161286406219
step: 400, loss: 0.004758063703775406
step: 410, loss: 0.003815303323790431
step: 420, loss: 0.00014928368909750134
step: 430, loss: 0.03176030516624451
step: 440, loss: 7.322348392335698e-05
step: 450, loss: 0.007098865229636431
step: 460, loss: 0.026609573513269424
step: 470, loss: 5.928441169089638e-05
step: 480, loss: 4.0433424146613106e-05

```
step: 490, loss: 0.02519090846180916
step: 500, loss: 0.0022861596662551165
step: 510, loss: 0.0017599723068997264
step: 520, loss: 0.009394537657499313
step: 530, loss: 0.15243938565254211
/usr/local/lib/python3.7/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=0.9407407407407408, f1=0.9394495412844037,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012511645909398794
step: 10, loss: 0.0018410427728667855
step: 20, loss: 0.0009698260109871626
step: 30, loss: 0.012869482859969139
step: 40, loss: 0.0033758245408535004
step: 50, loss: 0.003966032061725855
step: 60, loss: 0.013925706967711449
step: 70, loss: 0.00025475784786976874
step: 80, loss: 0.005073506850749254
step: 90, loss: 0.0070862979628145695
step: 100, loss: 0.0008982808794826269
step: 110, loss: 0.0026571720372885466
step: 120, loss: 0.08144856989383698
step: 130, loss: 0.0002650766109582037
step: 140, loss: 0.0020946357399225235
step: 150, loss: 0.0077498252503573895
step: 160, loss: 0.023664314299821854
step: 170, loss: 0.0034636908676475286
step: 180, loss: 0.0014163508312776685
step: 190, loss: 0.0002013984340010211
step: 200, loss: 0.004886111710220575
step: 210, loss: 0.00941146444529295
step: 220, loss: 0.002678218763321638
step: 230, loss: 0.06943763047456741
step: 240, loss: 0.0016322557348757982
step: 250, loss: 0.0005355554167181253
step: 260, loss: 0.004128619562834501
step: 270, loss: 0.00013858708553016186
step: 280, loss: 0.0006304687121883035
step: 290, loss: 0.0009705129195936024
step: 300, loss: 0.002703799633309245
step: 310, loss: 0.0239659883081913
step: 320, loss: 0.055129650980234146
step: 330, loss: 0.004746661055833101
step: 340, loss: 0.0018136565340682864
```

```

step: 350, loss: 0.0008561498834751546
step: 360, loss: 0.00013869788381271064
step: 370, loss: 0.0025098081678152084
step: 380, loss: 0.03033541329205036
step: 390, loss: 0.0002535638923291117
step: 400, loss: 0.00036528115742839873
step: 410, loss: 0.0012832414358854294
step: 420, loss: 0.0006113097770139575
step: 430, loss: 0.0018583715427666903
step: 440, loss: 0.0001872796128736809
step: 450, loss: 0.0010270837228745222
step: 460, loss: 0.015527365729212761
step: 470, loss: 0.0043870932422578335
step: 480, loss: 0.0008395116892643273
step: 490, loss: 0.020415084436535835
step: 500, loss: 0.0003713170881383121
step: 510, loss: 0.0002644560590852052
step: 520, loss: 0.0027984946500509977
step: 530, loss: 0.008947985246777534
/usr/local/lib/python3.7/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=0.9468331021729081, f1=0.9406896551724139,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.129491859814152e-05
step: 10, loss: 0.0028599353972822428
step: 20, loss: 0.00025183873367495835
step: 30, loss: 0.0005771557916887105
step: 40, loss: 4.517481283983216e-05
step: 50, loss: 5.984805466141552e-05
step: 60, loss: 0.0003651083679869771
step: 70, loss: 0.0012997087324038148
step: 80, loss: 0.001203400082886219
step: 90, loss: 0.0007135743508115411
step: 100, loss: 0.006839582230895758
step: 110, loss: 0.0007944925455376506
step: 120, loss: 0.002192085376009345
step: 130, loss: 0.0031044832430779934
step: 140, loss: 0.0015595054719597101
step: 150, loss: 0.0031502740457654
step: 160, loss: 0.001208788831718266
step: 170, loss: 0.003875638358294964
step: 180, loss: 0.0008071581833064556
step: 190, loss: 0.0005055352812632918
step: 200, loss: 0.0007827008957974613

```

```

step: 210, loss: 0.0005269284010864794
step: 220, loss: 0.004200655035674572
step: 230, loss: 0.0010597821092233062
step: 240, loss: 0.0014860149240121245
step: 250, loss: 0.03323015198111534
step: 260, loss: 0.029082899913191795
step: 270, loss: 0.003539895173162222
step: 280, loss: 0.005252073053270578
step: 290, loss: 0.0038017428014427423
step: 300, loss: 0.013089069165289402
step: 310, loss: 0.0015913064125925303
step: 320, loss: 7.126209675334394e-05
step: 330, loss: 3.1399944418808445e-05
step: 340, loss: 0.012929671443998814
step: 350, loss: 0.00041444774251431227
step: 360, loss: 0.0031718506943434477
step: 370, loss: 0.0015635319286957383
step: 380, loss: 0.0016826495993882418
step: 390, loss: 0.0028361016884446144
step: 400, loss: 4.984647239325568e-05
step: 410, loss: 0.0006381837883964181
step: 420, loss: 0.0029353906866163015
step: 430, loss: 0.002645349595695734
step: 440, loss: 0.00014463224215433002
step: 450, loss: 0.001953497529029846
step: 460, loss: 9.527475049253553e-05
step: 470, loss: 0.0006134920404292643
step: 480, loss: 0.0002464063873048872
step: 490, loss: 0.005606933031231165
step: 500, loss: 0.0018546974752098322
step: 510, loss: 0.0012365153525024652
step: 520, loss: 0.00027225996018387377
step: 530, loss: 0.02801484428346157
/usr/local/lib/python3.7/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=0.9425287356321841, f1=0.9399358092618065,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004622664302587509
step: 10, loss: 0.001639554975554347
step: 20, loss: 0.0019140229560434818
step: 30, loss: 2.3469030566047877e-05
step: 40, loss: 0.0006422038422897458
step: 50, loss: 0.0001390898396493867
step: 60, loss: 0.00015852091019041836

```

step: 70, loss: 0.17298537492752075
step: 80, loss: 8.820345101412386e-05
step: 90, loss: 0.001853913301602006
step: 100, loss: 0.32258638739585876
step: 110, loss: 0.004405909217894077
step: 120, loss: 0.0034233778715133667
step: 130, loss: 0.004768860060721636
step: 140, loss: 0.00014075699436943978
step: 150, loss: 0.0004067036497872323
step: 160, loss: 0.0021316860802471638
step: 170, loss: 0.0006086903158575296
step: 180, loss: 0.0002109039924107492
step: 190, loss: 0.0009252864983864129
step: 200, loss: 0.0018693895544856787
step: 210, loss: 0.0005143804592080414
step: 220, loss: 0.0026840665377676487
step: 230, loss: 0.0010376524878665805
step: 240, loss: 0.0013700282434001565
step: 250, loss: 6.100461905589327e-05
step: 260, loss: 0.0005405879346653819
step: 270, loss: 9.711955499369651e-05
step: 280, loss: 0.0033475186210125685
step: 290, loss: 0.0033407495357096195
step: 300, loss: 0.011944049037992954
step: 310, loss: 0.0008299993351101875
step: 320, loss: 0.006149237975478172
step: 330, loss: 0.03876996785402298
step: 340, loss: 0.00043917750008404255
step: 350, loss: 0.0004165095160715282
step: 360, loss: 0.0006244293763302267
step: 370, loss: 0.001216345583088696
step: 380, loss: 0.00019991632143501192
step: 390, loss: 0.0006290408200584352
step: 400, loss: 6.528422818519175e-05
step: 410, loss: 0.000559281324967742
step: 420, loss: 0.007952719926834106
step: 430, loss: 0.0062842415645718575
step: 440, loss: 0.000666198437102139
step: 450, loss: 0.00407251063734293
step: 460, loss: 4.4428001274354756e-05
step: 470, loss: 0.0010464175138622522
step: 480, loss: 0.0007324417820200324
step: 490, loss: 0.00015076829004101455
step: 500, loss: 7.872544665588066e-05
step: 510, loss: 0.057788971811532974
step: 520, loss: 0.001688378513790667
step: 530, loss: 0.001179031445644796
/usr/local/lib/python3.7/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=0.9487648673376029, f1=0.9457858769931662,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001420232729287818
step: 10, loss: 0.0015574427088722587
step: 20, loss: 0.00010855622531380504
step: 30, loss: 3.3884207368828356e-05
step: 40, loss: 0.004556823056191206
step: 50, loss: 0.002252815291285515
step: 60, loss: 2.724626028793864e-05
step: 70, loss: 0.0012604743242263794
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.0007655859808437526
step: 90, loss: 0.0001958157808985561
step: 100, loss: 0.0030006933957338333
step: 110, loss: 0.0002522230497561395
step: 120, loss: 0.0010587701108306646
step: 130, loss: 0.00011037817603209987
step: 140, loss: 0.011892483569681644
step: 150, loss: 0.001276044175028801
step: 160, loss: 0.0004103960527572781
step: 170, loss: 0.0019748739432543516
step: 180, loss: 0.0005825518164783716
step: 190, loss: 0.00041007468826137483
step: 200, loss: 2.73608930001501e-05
step: 210, loss: 0.0018666342366486788
step: 220, loss: 0.000115916998765897
step: 230, loss: 0.0024783117696642876
step: 240, loss: 0.0008311852579936385
step: 250, loss: 5.0627342716325074e-05
step: 260, loss: 0.061407916247844696
step: 270, loss: 3.3221320336451754e-05
step: 280, loss: 0.0010374155826866627
step: 290, loss: 3.375595770194195e-05
step: 300, loss: 2.009396666835528e-05
step: 310, loss: 7.305815233848989e-05
step: 320, loss: 0.00015797509695403278
step: 330, loss: 0.005856194067746401
step: 340, loss: 0.010663222521543503
step: 350, loss: 0.0018575915601104498
step: 360, loss: 0.09124357998371124
step: 370, loss: 0.0006091097020544112
step: 380, loss: 2.5796634872676805e-05
step: 390, loss: 7.887057290645316e-05

```



```

step: 400, loss: 7.246114546433091e-05
step: 410, loss: 8.168644126271829e-05
step: 420, loss: 0.000508182100020349
step: 430, loss: 0.0001971853052964434
step: 440, loss: 0.0001259334821952507
step: 450, loss: 0.020665045827627182
step: 460, loss: 0.0014968763571232557
step: 470, loss: 0.004719669930636883
step: 480, loss: 1.7862335880636238e-05
step: 490, loss: 8.608513599028811e-05
step: 500, loss: 2.1807187295053154e-05
step: 510, loss: 0.0001703365269349888
step: 520, loss: 2.552485602791421e-05
step: 530, loss: 3.0660812626592815e-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=0.9485396383866481, f1=0.9424758398527383,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004342400352470577
step: 10, loss: 0.0007217947859317064
step: 20, loss: 2.2775071556679904e-05
step: 30, loss: 0.03048732504248619
step: 40, loss: 0.00035212640068493783
step: 50, loss: 0.0004678008845075965
step: 60, loss: 0.0008293272694572806
step: 70, loss: 0.00022184144472703338
step: 80, loss: 0.0005145965842530131
step: 90, loss: 2.812348975567147e-05
step: 100, loss: 2.2060628907638602e-05
step: 110, loss: 0.00016255021910183132
step: 120, loss: 5.381739538279362e-05
step: 130, loss: 0.0003800346457865089
step: 140, loss: 0.0023837625049054623
step: 150, loss: 0.0034875082783401012
step: 160, loss: 0.001247349544428289
step: 170, loss: 0.00016988349671009928
step: 180, loss: 0.00015561329200863838
step: 190, loss: 0.0020571162458509207
step: 200, loss: 0.00012843866716139019
step: 210, loss: 0.01575685851275921
step: 220, loss: 1.6472922652610578e-05
step: 230, loss: 0.000333186995703727
step: 240, loss: 0.0020844340324401855
step: 250, loss: 0.001919322065077722

```

```

step: 260, loss: 0.00313810957595706
step: 270, loss: 0.0025389075744897127
step: 280, loss: 5.7821998780127615e-05
step: 290, loss: 7.75705193518661e-05
step: 300, loss: 2.0611647414625622e-05
step: 310, loss: 3.3670828997856006e-05
step: 320, loss: 5.32373487658333e-05
step: 330, loss: 0.0016231213230639696
step: 340, loss: 0.0006051517557352781
step: 350, loss: 0.00013521935034077615
step: 360, loss: 4.921403888147324e-05
step: 370, loss: 0.0003841146535705775
step: 380, loss: 0.0002571730292402208
step: 390, loss: 0.00433031702414155
step: 400, loss: 0.0024363910779356956
step: 410, loss: 1.8857113900594413e-05
step: 420, loss: 0.0003463909379206598
step: 430, loss: 0.23893171548843384
step: 440, loss: 0.009395492263138294
step: 450, loss: 2.864609814423602e-05
step: 460, loss: 0.023486211895942688
step: 470, loss: 0.00017045668209902942
step: 480, loss: 0.00042917224345728755
step: 490, loss: 0.0004770659434143454
step: 500, loss: 0.002746665384620428
step: 510, loss: 0.01524451095610857
step: 520, loss: 0.0005102226859889925
step: 530, loss: 0.0017612582305446267
/usr/local/lib/python3.7/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=0.9452181987000929, f1=0.94547134935305,
best_f1=0.9466911764705882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016812178364489228
step: 10, loss: 0.0014646017225459218
step: 20, loss: 0.0011518137762323022
step: 30, loss: 0.004499153699725866
step: 40, loss: 0.0001411809353157878
step: 50, loss: 0.0034377831034362316
step: 60, loss: 0.010872084647417068
step: 70, loss: 0.00012266798876225948
step: 80, loss: 0.00013895878510084003
step: 90, loss: 0.0003151529817841947
step: 100, loss: 0.00018544877821113914
step: 110, loss: 0.0029468899592757225

```

```
step: 120, loss: 0.000158568233018741
step: 130, loss: 0.00909784622490406
step: 140, loss: 6.663425301667303e-05
step: 150, loss: 0.00015085334598552436
step: 160, loss: 0.00011335467570461333
step: 170, loss: 0.0005695211584679782
step: 180, loss: 0.0025327452458441257
step: 190, loss: 0.00289501273073256
step: 200, loss: 0.00341470493003726
step: 210, loss: 0.00017569582269061357
step: 220, loss: 0.0006023910245858133
step: 230, loss: 0.003134890226647258
step: 240, loss: 0.025503678247332573
step: 250, loss: 0.00014024948177393526
step: 260, loss: 0.0004433748545125127
step: 270, loss: 0.000502136186696589
step: 280, loss: 4.640388215193525e-05
step: 290, loss: 6.0776390455430374e-05
step: 300, loss: 0.0003783343418035656
step: 310, loss: 0.00952970515936613
step: 320, loss: 7.018547330517322e-05
step: 330, loss: 6.520382885355502e-05
step: 340, loss: 0.001289075706154108
step: 350, loss: 0.004455302841961384
step: 360, loss: 0.0026052219327539206
step: 370, loss: 0.003912735264748335
step: 380, loss: 0.00042789377039298415
step: 390, loss: 0.0637037381529808
step: 400, loss: 0.0027492432855069637
step: 410, loss: 0.000858452869579196
step: 420, loss: 0.00011667457147268578
step: 430, loss: 5.352950029191561e-05
step: 440, loss: 0.0003363971773069352
step: 450, loss: 0.008881657384335995
step: 460, loss: 0.005197683349251747
step: 470, loss: 3.284795820945874e-05
step: 480, loss: 0.004953771363943815
step: 490, loss: 0.0019998704083263874
step: 500, loss: 0.002305199857801199
step: 510, loss: 0.0005047296290285885
step: 520, loss: 8.7915534095373e-05
step: 530, loss: 0.00029713514959439635
/usr/local/lib/python3.7/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=0.9472222222222223, f1=0.9455719557195572,
best_f1=0.9466911764705882
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 [00:21, 266.55it/s]

load_f1 = 0.9474654377880185

real_f1 = 0.9483870967741935

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

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

4398it [00:21, 208.60it/s]