

HLow_10_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/Snippest_public
      %cd Snippest_public
      !pip install -r requirements.txt
      %cd ..

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 20.94 MiB/s, done.
Resolving deltas: 100% (142/142), done.
Checking out files: 100% (128/128), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.4 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 49.5 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 53.7 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 83.4 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 53.3 MB/s
Requirement already satisfied: sklearn==0.0 in
```

```

/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
  | 6.4 MB 52.8 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 20.21 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 5.2 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 76.1 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
  | 753.4 MB 6.8 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
  | 56 kB 6.6 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
  | 884 kB 70.5 MB/s
Collecting jsonlines==1.2.0

```

```

    Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)
Collecting nltk==3.4.5
    Downloading nltk-3.4.5.zip (1.5 MB)
      |                               | 1.5 MB 59.0 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: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
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 57.4 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 blis<0.5.0,>=0.4.0
    Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
      |                               | 3.7 MB 47.3 MB/s
Collecting plac<1.2.0,>=0.9.6
    Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting srsly<1.1.0,>=0.1.0
    Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
      |                               | 184 kB 78.4 MB/s
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: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: 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 catalogue<1.1.0,>=0.0.7
    Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
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: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting sacremoses

```

```

Downloading sacremoses-0.0.53.tar.gz (880 kB)
|                                     | 880 kB 67.7 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    |                                     | 3.0 MB 36.1 MB/s
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: 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: 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: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
    Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449923
    sha256=77c56ab5d57e3f23baf7f96d87648cdfd4d489513dfe127a07a5f89a9f5d09c3
    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=4bc9cf22d7214059af58524fd30b1f83419db942eb52e8d91d908912c836b5ac
  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 | 21.56 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-cgpaf6as
Created temporary directory: /tmp/pip-req-tracker-p53exglw
Initialized build tracking at /tmp/pip-req-tracker-p53exglw
Created build tracker: /tmp/pip-req-tracker-p53exglw
Entered build tracker: /tmp/pip-req-tracker-p53exglw
```

```

Created temporary directory: /tmp/pip-install-kvij5hiy
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-b6_joiow
  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-p53exglw'
  Running setup.py (path:/tmp/pip-req-build-b6_joiow/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-_u4wj9x8
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-_u4wj9x8/apex.egg-info
  writing /tmp/pip-pip-egg-info-_u4wj9x8/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-_u4wj9x8/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-_u4wj9x8/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-_u4wj9x8/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-_u4wj9x8/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-b6_joiow 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-p53exglw'
Created temporary directory: /tmp/pip-unpack-c77mv28_
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-riofnjgq
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-riofnjgq
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-b6_joiow/setup.py''';

```



```
__file__='''/tmp/pip-req-build-b6_joiow/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-riofnjgq
```

```
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_parallel_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_parallel_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-riofnjgq/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=74dbd0202546c30c1d16442496a48052d096900a27c87468e9b8f2eec6cfbc3f
Stored in directory: /tmp/pip-ephem-wheel-cache-
cgpaf6as/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-p53exglw'
/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 5.1 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 35.9 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 55.9 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting boto3==1.27.48
  Downloading boto3-1.27.48-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 12.1 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 4.0 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: numpy in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1) (1.19.2)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

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

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

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 | 29.33 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/HLow_10_1_2/configs.json /content/ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm 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, 414kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 678kB/s]

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

Downloading: 100% 501M/501M [00:07<00:00, 63.6MB/s]

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

Defaults for this optimization level are:

enabled : True

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4606929421424866
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=0.2692307692307693, f1=0.2666666666666667,
best_f1=0.2666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3912649154663086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6285714285714286, f1=0.6486486486486486,
best_f1=0.6486486486486486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39043593406677246
/usr/local/lib/python3.7/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=0.4, f1=0.2777777777777778, best_f1=0.6486486486486486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2605190575122833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.896551724137931, f1=0.7878787878787878,
best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09877368062734604
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 5: dev_f1=0.8666666666666666, f1=0.8484848484848484,
best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09889082610607147
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.7741935483870968,
best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22774793207645416
/usr/local/lib/python3.7/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=0.8888888888888889, f1=0.9032258064516129,
best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08646609634160995
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.8750000000000001,
best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01406009215861559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9285714285714286, f1=0.9032258064516129,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08510123938322067
/usr/local/lib/python3.7/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=0.8750000000000001, f1=0.8750000000000001,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019512301310896873
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009546658955514431
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00740770110860467
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006892245728522539
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006899528671056032
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129
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, 111309.90it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.896551724137931
real_f1 = 0.8666666666666666
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 [00:29, 147.78it/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.")
```

step: 0, loss: 0.5670790672302246

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

step: 10, loss: 0.47089719772338867

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

step: 20, loss: 0.44108158349990845

step: 30, loss: 0.26246577501296997

step: 40, loss: 0.12632408738136292

step: 50, loss: 0.08186665922403336

```

step: 60, loss: 0.16771619021892548
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.09527211636304855
step: 80, loss: 0.025055095553398132
step: 90, loss: 0.08267748355865479
step: 100, loss: 0.09157966822385788
step: 110, loss: 0.016116971150040627
step: 120, loss: 0.024409513920545578
step: 130, loss: 0.0058412631042301655
step: 140, loss: 0.01710767112672329
step: 150, loss: 0.08170495182275772
step: 160, loss: 0.0027773373294621706
step: 170, loss: 0.06369596719741821
step: 180, loss: 0.050697147846221924
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 190, loss: 0.0016002345364540815
step: 200, loss: 0.036579664796590805
step: 210, loss: 0.04949736222624779
step: 220, loss: 0.007746001239866018
step: 230, loss: 0.0035245458129793406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9723756906077348, f1=0.9700996677740864,
best_f1=0.9700996677740864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0045689186081290245
step: 10, loss: 0.10923172533512115
step: 20, loss: 0.01595505326986313
step: 30, loss: 0.011596428230404854
step: 40, loss: 0.041888825595378876
step: 50, loss: 0.0066221910528838634
step: 60, loss: 0.0019316693069413304
step: 70, loss: 0.00507716927677393
step: 80, loss: 0.007684351410716772
step: 90, loss: 0.11190406233072281
step: 100, loss: 0.004845107439905405
step: 110, loss: 0.00722459377720952
step: 120, loss: 0.004988746251910925
step: 130, loss: 0.001695446902886033
step: 140, loss: 0.004334024153649807
step: 150, loss: 0.13308893144130707

```

```

step: 160, loss: 0.003358755260705948
step: 170, loss: 0.0005955794476903975
step: 180, loss: 0.010902724228799343
step: 190, loss: 0.06312959641218185
step: 200, loss: 0.0035237742122262716
step: 210, loss: 0.0024899544660001993
step: 220, loss: 0.0010872259736061096
step: 230, loss: 0.0009839278645813465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.984090909090909, f1=0.9794988610478361,
best_f1=0.9794988610478361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0067986552603542805
step: 10, loss: 0.00572030758485198
step: 20, loss: 0.029565053060650826
step: 30, loss: 0.006930357776582241
step: 40, loss: 0.018646955490112305
step: 50, loss: 0.262594997882843
step: 60, loss: 0.01251621451228857
step: 70, loss: 0.003690196666866541
step: 80, loss: 0.003966622985899448
step: 90, loss: 0.004406317602843046
step: 100, loss: 0.0051814718171954155
step: 110, loss: 0.00795313436537981
step: 120, loss: 0.0017009121365845203
step: 130, loss: 0.0031700595282018185
step: 140, loss: 0.001538179349154234
step: 150, loss: 0.08788007497787476
step: 160, loss: 0.00898483395576477
step: 170, loss: 0.0007301673176698387
step: 180, loss: 0.0015561498003080487
step: 190, loss: 0.008695363998413086
step: 200, loss: 0.00308287818916142
step: 210, loss: 0.003471909323707223
step: 220, loss: 0.004885638132691383
step: 230, loss: 0.005952429957687855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9921259842519685, f1=0.9797752808988766,

best_f1=0.9797752808988766

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

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

step: 0, loss: 0.0017089968314394355

step: 10, loss: 0.001121927984058857

step: 20, loss: 0.001019461895339191

step: 30, loss: 0.0011016756761819124

step: 40, loss: 0.0025691490154713392

step: 50, loss: 0.0021565810311585665

step: 60, loss: 0.0023122343700379133

step: 70, loss: 0.0008376328623853624

step: 80, loss: 0.00039002037374302745

step: 90, loss: 0.000806903000921011

step: 100, loss: 0.0011800737120211124

step: 110, loss: 0.0005588948843069375

step: 120, loss: 0.0018434006487950683

step: 130, loss: 0.043145183473825455

step: 140, loss: 0.004060812760144472

step: 150, loss: 0.0007897233590483665

step: 160, loss: 0.003098632674664259

step: 170, loss: 0.005031007342040539

step: 180, loss: 0.15912356972694397

step: 190, loss: 0.0009719772497192025

step: 200, loss: 0.004085164982825518

step: 210, loss: 0.0008736054296605289

step: 220, loss: 0.0004683480365201831

step: 230, loss: 0.02290109358727932

/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9799107142857142,

best_f1=0.9797752808988766

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

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

step: 0, loss: 0.0005349192651920021

step: 10, loss: 0.0009113380219787359

step: 20, loss: 0.012175030075013638

step: 30, loss: 0.0006277511129155755

step: 40, loss: 0.011382979340851307

step: 50, loss: 0.0025254483334720135

step: 60, loss: 0.0005494877113960683

step: 70, loss: 0.0007922116783447564

```

step: 80, loss: 0.02293543703854084
step: 90, loss: 0.03879234567284584
step: 100, loss: 0.0026374836452305317
step: 110, loss: 0.002672688802704215
step: 120, loss: 0.00033230811823159456
step: 130, loss: 0.00026641422300599515
step: 140, loss: 0.0008736575255170465
step: 150, loss: 0.039412446320056915
step: 160, loss: 0.0002466639271005988
step: 170, loss: 0.0006751324981451035
step: 180, loss: 0.0005942051066085696
step: 190, loss: 0.004653926473110914
step: 200, loss: 0.017398307099938393
step: 210, loss: 0.0051793004386126995
step: 220, loss: 0.0011771017452701926
step: 230, loss: 0.0027123780455440283
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9842342342342343,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018949776131194085
step: 10, loss: 0.0012989819515496492
step: 20, loss: 0.0009929215302690864
step: 30, loss: 0.000590314797591418
step: 40, loss: 0.0007158025400713086
step: 50, loss: 0.00038635157397948205
step: 60, loss: 0.004457143135368824
step: 70, loss: 0.2551650404930115
step: 80, loss: 0.0005358346970751882
step: 90, loss: 0.001272753463126719
step: 100, loss: 0.0006156771560199559
step: 110, loss: 0.01021910086274147
step: 120, loss: 0.0009356348891742527
step: 130, loss: 0.0016731226351112127
step: 140, loss: 0.0021532243117690086
step: 150, loss: 0.00021392732742242515
step: 160, loss: 0.00859943125396967
step: 170, loss: 0.00027232366846874356
step: 180, loss: 0.00020758395839948207
step: 190, loss: 0.00014944199938327074
step: 200, loss: 0.0009837474208325148
step: 210, loss: 0.000289840972982347
step: 220, loss: 0.028227653354406357
step: 230, loss: 0.022767364978790283

```

```

/usr/local/lib/python3.7/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=0.9864864864864865, f1=0.9831649831649831,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003265458159148693
step: 10, loss: 0.0017618455458432436
step: 20, loss: 0.002640157239511609
step: 30, loss: 0.0012143636122345924
step: 40, loss: 0.0018282791133970022
step: 50, loss: 0.0006309236632660031
step: 60, loss: 0.0005192090757191181
step: 70, loss: 0.0012817785609513521
step: 80, loss: 0.0004118015931453556
step: 90, loss: 0.00942891463637352
step: 100, loss: 0.0002421817189315334
step: 110, loss: 0.00026270459056831896
step: 120, loss: 0.0003013556415680796
step: 130, loss: 0.00018169707618653774
step: 140, loss: 0.00011216787970624864
step: 150, loss: 0.00123275734949857
step: 160, loss: 0.07840240001678467
step: 170, loss: 0.003913154825568199
step: 180, loss: 0.0012932345271110535
step: 190, loss: 0.00017619514255784452
step: 200, loss: 0.00031373731326311827
step: 210, loss: 0.00045918775140307844
step: 220, loss: 0.0003573840076569468
step: 230, loss: 0.00018002941214945167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9932584269662922, f1=0.9854423292273236,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001727151102386415
step: 10, loss: 0.0004024270747322589
step: 20, loss: 0.0003342995187267661
step: 30, loss: 0.00034750605118460953

```



```

step: 40, loss: 0.0021672253496944904
step: 50, loss: 0.00041575226350687444
step: 60, loss: 0.0002982866717502475
step: 70, loss: 0.00012760995014104992
step: 80, loss: 0.017446065321564674
step: 90, loss: 0.00032525291317142546
step: 100, loss: 0.0002354671887587756
step: 110, loss: 0.00022635774803347886
step: 120, loss: 0.000490992795675993
step: 130, loss: 0.0048161642625927925
step: 140, loss: 0.000483944546431303
step: 150, loss: 0.06543421745300293
step: 160, loss: 0.0017595603130757809
step: 170, loss: 0.006029743235558271
step: 180, loss: 0.00044076089398004115
step: 190, loss: 0.0010808843653649092
step: 200, loss: 0.0002514664956834167
step: 210, loss: 0.0011656078277155757
step: 220, loss: 0.0004266418400220573
step: 230, loss: 0.0003530179092194885
/usr/local/lib/python3.7/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=0.9921612541993281, f1=0.9844444444444443,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030258670449256897
step: 10, loss: 0.0006392306531779468
step: 20, loss: 0.0002912808849941939
step: 30, loss: 0.0008151168585754931
step: 40, loss: 0.00016390992095693946
step: 50, loss: 0.000495576299726963
step: 60, loss: 0.00012724872794933617
step: 70, loss: 0.028429457917809486
step: 80, loss: 0.00038726176717318594
step: 90, loss: 0.031742800027132034
step: 100, loss: 0.00012529171362984926
step: 110, loss: 0.00011421323870308697
step: 120, loss: 0.028987377882003784
step: 130, loss: 0.00045014554052613676
step: 140, loss: 0.0007394712883979082
step: 150, loss: 0.0006087086512707174
step: 160, loss: 0.0019812355749309063
step: 170, loss: 0.0002015204227063805
step: 180, loss: 0.0010457697790116072
step: 190, loss: 0.00015989845269359648

```

```

step: 200, loss: 0.00021927007765043527
step: 210, loss: 0.03398658335208893
step: 220, loss: 0.00033755795448087156
step: 230, loss: 0.00017174922686535865
/usr/local/lib/python3.7/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=0.9921259842519685, f1=0.9887640449438202,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030672617140226066
step: 10, loss: 0.0005356633337214589
step: 20, loss: 0.0005291011184453964
step: 30, loss: 0.0002661280450411141
step: 40, loss: 0.003050596918910742
step: 50, loss: 0.0002135286049451679
step: 60, loss: 0.0004019818443339318
step: 70, loss: 0.0005812803865410388
step: 80, loss: 0.0002462312113493681
step: 90, loss: 0.00020812859293073416
step: 100, loss: 0.00012710073497146368
step: 110, loss: 0.000692566973157227
step: 120, loss: 0.00011941859702346846
step: 130, loss: 0.0002368529821978882
step: 140, loss: 0.00015847165195737034
step: 150, loss: 0.00041587790474295616
step: 160, loss: 7.871663547120988e-05
step: 170, loss: 0.00015321550017688423
step: 180, loss: 0.0005229049711488187
step: 190, loss: 0.0003636165929492563
step: 200, loss: 0.0006354870274662971
step: 210, loss: 0.00035840042983181775
step: 220, loss: 0.00014162450679577887
step: 230, loss: 0.00015229158452712
/usr/local/lib/python3.7/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=0.9921259842519685, f1=0.988814317673378,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030633850838057697
step: 10, loss: 0.0005340678617358208
step: 20, loss: 0.0002691383706405759
step: 30, loss: 0.00019022281048819423

```

```

step: 40, loss: 3.992297570221126e-05
step: 50, loss: 0.00010385439964011312
step: 60, loss: 0.011492766439914703
step: 70, loss: 0.0001386203512083739
step: 80, loss: 0.00031252173357643187
step: 90, loss: 0.06237746402621269
step: 100, loss: 0.0005625477060675621
step: 110, loss: 0.0004632473865058273
step: 120, loss: 0.000561052467674017
step: 130, loss: 0.00029678959981538355
step: 140, loss: 0.0036007266025990248
step: 150, loss: 0.00034007098292931914
step: 160, loss: 0.05053159222006798
step: 170, loss: 0.002997914096340537
step: 180, loss: 0.0006357739330269396
step: 190, loss: 0.00032557127997279167
step: 200, loss: 0.002706692088395357
step: 210, loss: 0.00023757631424814463
step: 220, loss: 0.0012585988733917475
step: 230, loss: 0.00017503740673419088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9921436588103255, f1=0.9833147942157954,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029212841764092445
step: 10, loss: 0.00011897948570549488
step: 20, loss: 0.011084440164268017
step: 30, loss: 0.04038737714290619
step: 40, loss: 0.0005525014130398631
step: 50, loss: 0.00048518087714910507
step: 60, loss: 0.00036932885996066034
step: 70, loss: 0.0003071457613259554
step: 80, loss: 0.00010077365732286125
step: 90, loss: 0.0018305721459910274
step: 100, loss: 0.00017407556879334152
step: 110, loss: 0.0001331573585048318
step: 120, loss: 0.0001424921356374398
step: 130, loss: 0.00013544867397285998
step: 140, loss: 0.0002577520499471575
step: 150, loss: 0.00025831168750301003
step: 160, loss: 0.00032885337714105844
step: 170, loss: 0.0001801802427507937
step: 180, loss: 0.002684551989659667
step: 190, loss: 0.0004949290887452662

```

```

step: 200, loss: 0.00017661054152995348
step: 210, loss: 0.0003454624966252595
step: 220, loss: 0.0178071316331625
step: 230, loss: 0.0016276435926556587
/usr/local/lib/python3.7/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=0.9921436588103255, f1=0.9844097995545658,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011385896243155003
step: 10, loss: 0.00032713834661990404
step: 20, loss: 0.0005103253643028438
step: 30, loss: 0.0010391048854216933
step: 40, loss: 0.0009102019248530269
step: 50, loss: 0.0020247502252459526
step: 60, loss: 0.0005165262264199555
step: 70, loss: 0.0015326502034440637
step: 80, loss: 0.0003729049058165401
step: 90, loss: 0.00028700451366603374
step: 100, loss: 0.0007539456128142774
step: 110, loss: 0.0003764186694752425
step: 120, loss: 0.00023055316705722362
step: 130, loss: 0.00018650630954653025
step: 140, loss: 0.0001059343121596612
step: 150, loss: 0.0008763388614170253
step: 160, loss: 0.004652957431972027
step: 170, loss: 0.00017022801330313087
step: 180, loss: 0.014061097986996174
step: 190, loss: 0.00025810382794588804
step: 200, loss: 4.450060077942908e-05
step: 210, loss: 0.00016558857169002295
step: 220, loss: 0.0001243510196218267
step: 230, loss: 0.0001456415484426543
/usr/local/lib/python3.7/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=0.9921436588103255, f1=0.9843749999999999,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001159877356258221
step: 10, loss: 8.866605639923364e-05
step: 20, loss: 0.00014894669584464282
step: 30, loss: 0.0001561577955726534

```

```

step: 40, loss: 0.00012759592209476978
step: 50, loss: 7.117416680557653e-05
step: 60, loss: 9.019980643643066e-05
step: 70, loss: 0.00011647478822851554
step: 80, loss: 6.004859824315645e-05
step: 90, loss: 0.000322470732498914
step: 100, loss: 0.00030864731525070965
step: 110, loss: 0.0004052604781463742
step: 120, loss: 4.824290590477176e-05
step: 130, loss: 0.0005991163197904825
step: 140, loss: 0.00019305618479847908
step: 150, loss: 0.00011579024430830032
step: 160, loss: 0.0003232496965210885
step: 170, loss: 0.00016066536772996187
step: 180, loss: 0.0007617890369147062
step: 190, loss: 0.00023096903169061989
step: 200, loss: 0.0012012351071462035
step: 210, loss: 0.0005942114512436092
step: 220, loss: 0.00032203554292209446
step: 230, loss: 9.550528920954093e-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=0.9921436588103255, f1=0.9844097995545658,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038917758502066135
step: 10, loss: 0.00019151403103023767
step: 20, loss: 0.0007945416728034616
step: 30, loss: 0.00033651210833340883
step: 40, loss: 0.00013335359108168632
step: 50, loss: 0.00015984600759111345
step: 60, loss: 0.04048025980591774
step: 70, loss: 0.00011499177344376221
step: 80, loss: 0.000113091038656421
step: 90, loss: 0.00010761293378891423
step: 100, loss: 7.082299998728558e-05
step: 110, loss: 0.0001338552829110995
step: 120, loss: 0.02561931498348713
step: 130, loss: 9.16865246836096e-05
step: 140, loss: 0.012527070939540863
step: 150, loss: 0.00020477989164646715
step: 160, loss: 0.01252803485840559
step: 170, loss: 5.754242738476023e-05
step: 180, loss: 0.0003778687969315797
step: 190, loss: 0.0014925767900422215

```

```

step: 200, loss: 0.004792036954313517
step: 210, loss: 0.033050425350666046
step: 220, loss: 0.00022253033239394426
step: 230, loss: 0.00013497244799509645
/usr/local/lib/python3.7/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=0.9921436588103255, f1=0.9844097995545658,
best_f1=0.9854423292273236
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 [00:16, 152.41it/s]
load_f1 = 0.9898534385569334
real_f1 = 0.9865168539325843
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, 138.51it/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                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6197522282600403
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: 0.41756075620651245
step: 20, loss: 0.2884119153022766
step: 30, loss: 0.37001165747642517
step: 40, loss: 0.3493230938911438
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.7366665601730347
step: 60, loss: 0.2058890163898468
step: 70, loss: 0.20747505128383636
step: 80, loss: 0.1096605509519577
step: 90, loss: 0.14316228032112122
step: 100, loss: 0.08184605836868286
step: 110, loss: 0.0901482030749321
step: 120, loss: 0.1265643835067749
step: 130, loss: 0.1582118272781372
step: 140, loss: 0.15751637518405914
step: 150, loss: 0.019107481464743614
step: 160, loss: 0.15467999875545502
step: 170, loss: 0.07283516973257065
step: 180, loss: 0.03912195563316345
step: 190, loss: 0.03462279587984085
step: 200, loss: 0.02493215538561344
step: 210, loss: 0.06324895471334457
step: 220, loss: 0.023654358461499214
step: 230, loss: 0.20444324612617493
step: 240, loss: 0.04878075420856476
step: 250, loss: 0.01818501576781273
step: 260, loss: 0.15647649765014648
step: 270, loss: 0.2772059738636017
step: 280, loss: 0.03144616633653641
step: 290, loss: 0.05163327232003212
step: 300, loss: 0.009710664860904217
step: 310, loss: 0.09087546169757843
step: 320, loss: 0.04082426428794861

```



```

step: 330, loss: 0.037834130227565765
step: 340, loss: 0.4436279535293579
step: 350, loss: 0.12518590688705444
step: 360, loss: 0.07904143631458282
step: 370, loss: 0.028834035620093346
step: 380, loss: 0.07736586779356003
step: 390, loss: 0.008490334264934063
step: 400, loss: 0.026179121807217598
step: 410, loss: 0.24157555401325226
step: 420, loss: 0.022753363475203514
step: 430, loss: 0.008972523733973503
step: 440, loss: 0.0218486450612545
step: 450, loss: 0.0509345643222332
step: 460, loss: 0.0065645440481603146
step: 470, loss: 0.020562028512358665
step: 480, loss: 0.2440144568681717
step: 490, loss: 0.13947738707065582
step: 500, loss: 0.006823567673563957
step: 510, loss: 0.015066049061715603
step: 520, loss: 0.045348577201366425
step: 530, loss: 0.024608053267002106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9455053563111318, f1=0.9476124246638852,
best_f1=0.9476124246638852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11333708465099335
step: 10, loss: 0.055147040635347366
step: 20, loss: 0.01808248832821846
step: 30, loss: 0.056394752115011215
step: 40, loss: 0.019220026209950447
step: 50, loss: 0.07420334219932556
step: 60, loss: 0.028141513466835022
step: 70, loss: 0.007702156435698271
step: 80, loss: 0.024332759901881218
step: 90, loss: 0.001821687095798552
step: 100, loss: 0.14856871962547302
step: 110, loss: 0.013336585834622383
step: 120, loss: 0.008677592501044273
step: 130, loss: 0.007924054749310017
step: 140, loss: 0.026011813431978226

```

```

step: 150, loss: 0.041307177394628525
step: 160, loss: 0.02890690416097641
step: 170, loss: 0.008790655992925167
step: 180, loss: 0.02487373538315296
step: 190, loss: 0.016182973980903625
step: 200, loss: 0.09544741362333298
step: 210, loss: 0.006013309117406607
step: 220, loss: 0.0008270477410405874
step: 230, loss: 0.09986905008554459
step: 240, loss: 0.06361079961061478
step: 250, loss: 0.06621187925338745
step: 260, loss: 0.017433734610676765
step: 270, loss: 0.007731188088655472
step: 280, loss: 0.10797761380672455
step: 290, loss: 0.040887512266635895
step: 300, loss: 0.017458172515034676
step: 310, loss: 0.04934681951999664
step: 320, loss: 0.058953575789928436
step: 330, loss: 0.05734558030962944
step: 340, loss: 0.007855340838432312
step: 350, loss: 0.003375562373548746
step: 360, loss: 0.05419351905584335
step: 370, loss: 0.002190870000049472
step: 380, loss: 0.1250172108411789
step: 390, loss: 0.004482632037252188
step: 400, loss: 0.007177895400673151
step: 410, loss: 0.006223029922693968
step: 420, loss: 0.021077753975987434
step: 430, loss: 0.18320685625076294
step: 440, loss: 0.006826520897448063
step: 450, loss: 0.0296097993850708
step: 460, loss: 0.030160481110215187
step: 470, loss: 0.013599122874438763
step: 480, loss: 0.06064896658062935
step: 490, loss: 0.02328963205218315
step: 500, loss: 0.0013542291708290577
step: 510, loss: 0.007998473010957241
step: 520, loss: 0.32346245646476746
step: 530, loss: 0.0708680972456932
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9504400185270958, f1=0.9549046954904695,
best_f1=0.9549046954904695

```

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

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

```
step: 0, loss: 0.0812847688794136  
step: 10, loss: 0.04349192976951599  
step: 20, loss: 0.009015989489853382  
step: 30, loss: 0.018741121515631676  
step: 40, loss: 0.04664833843708038  
step: 50, loss: 0.023599252104759216  
step: 60, loss: 0.0048384889960289  
step: 70, loss: 0.02768274024128914  
step: 80, loss: 0.0228207278996706  
step: 90, loss: 0.02542819269001484  
step: 100, loss: 0.03019702062010765  
step: 110, loss: 0.0350455678999424  
step: 120, loss: 0.22782368957996368  
step: 130, loss: 0.030705038458108902  
step: 140, loss: 0.01768733561038971  
step: 150, loss: 0.03154480829834938  
step: 160, loss: 0.111654132604599  
step: 170, loss: 0.004189169500023127  
step: 180, loss: 0.006365468259900808  
step: 190, loss: 0.027328142896294594  
step: 200, loss: 0.016565967351198196  
step: 210, loss: 0.01590389385819435  
step: 220, loss: 0.10664935410022736  
step: 230, loss: 0.054948460310697556  
step: 240, loss: 0.0746690034866333  
step: 250, loss: 0.11833063513040543  
step: 260, loss: 0.11455202102661133  
step: 270, loss: 0.0022133495658636093  
step: 280, loss: 0.03439915552735329  
step: 290, loss: 0.04807104170322418  
step: 300, loss: 0.14313030242919922  
step: 310, loss: 0.057625140994787216  
step: 320, loss: 0.02864054962992668  
step: 330, loss: 0.008586953394114971  
step: 340, loss: 0.007496899459511042  
step: 350, loss: 0.12390319257974625  
step: 360, loss: 0.0120805399492383  
step: 370, loss: 0.02658936195075512  
step: 380, loss: 0.011177041567862034  
step: 390, loss: 0.0049754963256418705  
step: 400, loss: 0.09492567181587219  
step: 410, loss: 0.04253162443637848  
step: 420, loss: 0.02185126207768917  
step: 430, loss: 0.015950500965118408  
step: 440, loss: 0.16656215488910675
```

```

step: 450, loss: 0.03450074419379234
step: 460, loss: 0.19661587476730347
step: 470, loss: 0.019798528403043747
step: 480, loss: 0.02024797908961773
step: 490, loss: 0.10128942877054214
step: 500, loss: 0.0369444340467453
step: 510, loss: 0.018198788166046143
step: 520, loss: 0.013709943741559982
step: 530, loss: 0.00845105666667223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9553903345724908, f1=0.957169459962756,
best_f1=0.957169459962756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01901591755449772
step: 10, loss: 0.004609879106283188
step: 20, loss: 0.1381228268146515
step: 30, loss: 0.14533059298992157
step: 40, loss: 0.009993769228458405
step: 50, loss: 0.0335395410656929
step: 60, loss: 0.011251246556639671
step: 70, loss: 0.0019951583817601204
step: 80, loss: 0.0014337475877255201
step: 90, loss: 0.06521543115377426
step: 100, loss: 0.00041488828719593585
step: 110, loss: 0.012279625050723553
step: 120, loss: 0.0012711695162579417
step: 130, loss: 0.11264701187610626
step: 140, loss: 0.043845344334840775
step: 150, loss: 0.0011213722173124552
step: 160, loss: 0.009721402078866959
step: 170, loss: 0.01643352583050728
step: 180, loss: 0.09182095527648926
step: 190, loss: 0.09146177768707275
step: 200, loss: 0.016461223363876343
step: 210, loss: 0.0004065603425260633
step: 220, loss: 0.0019113044254481792
step: 230, loss: 0.006284456700086594
step: 240, loss: 0.014831345528364182
step: 250, loss: 0.07217708975076675
step: 260, loss: 0.0009524640045128763

```

```

step: 270, loss: 0.06923214346170425
step: 280, loss: 0.008601237088441849
step: 290, loss: 0.032735515385866165
step: 300, loss: 0.0011500815162435174
step: 310, loss: 0.005286825355142355
step: 320, loss: 0.015235981903970242
step: 330, loss: 0.006994959432631731
step: 340, loss: 0.0019478583708405495
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 350, loss: 0.12737016379833221
step: 360, loss: 0.019209468737244606
step: 370, loss: 0.00581278745085001
step: 380, loss: 0.003803151659667492
step: 390, loss: 0.00031730293994769454
step: 400, loss: 0.006743682082742453
step: 410, loss: 0.001194600248709321
step: 420, loss: 0.015381026081740856
step: 430, loss: 0.03136785700917244
step: 440, loss: 0.002985595492646098
step: 450, loss: 0.03989849612116814
step: 460, loss: 0.007126485928893089
step: 470, loss: 0.007127605844289064
step: 480, loss: 0.001611255924217403
step: 490, loss: 0.0036703944206237793
step: 500, loss: 0.0674804076552391
step: 510, loss: 0.04732424393296242
step: 520, loss: 0.006466454826295376
step: 530, loss: 0.05004889890551567
/usr/local/lib/python3.7/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=0.9552376557452701, f1=0.9576816927322908,
best_f1=0.957169459962756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006349430186673999
step: 10, loss: 0.02924472652375698
step: 20, loss: 0.004939547274261713
step: 30, loss: 0.002929436042904854
step: 40, loss: 0.0011848704889416695
step: 50, loss: 0.03466646373271942
step: 60, loss: 0.01535507757216692
step: 70, loss: 0.0012690223520621657
step: 80, loss: 0.00030358927324414253
step: 90, loss: 0.09597376734018326
step: 100, loss: 0.009494178928434849
step: 110, loss: 0.0006363351712934673

```

```
step: 120, loss: 0.09647170454263687
step: 130, loss: 0.002175457775592804
step: 140, loss: 0.00043782684952020645
step: 150, loss: 0.0029363154899328947
step: 160, loss: 0.00335936457850039
step: 170, loss: 0.058227743953466415
step: 180, loss: 0.015976183116436005
step: 190, loss: 0.005410938989371061
step: 200, loss: 0.008042028173804283
step: 210, loss: 0.0010709782363846898
step: 220, loss: 0.0030241929925978184
step: 230, loss: 0.0008185736951418221
step: 240, loss: 0.0005390211008489132
step: 250, loss: 0.17794734239578247
step: 260, loss: 0.0008242272888310254
step: 270, loss: 0.02599356323480606
step: 280, loss: 0.0038158141542226076
step: 290, loss: 0.00047671739594079554
step: 300, loss: 0.0018940697191283107
step: 310, loss: 0.039044395089149475
step: 320, loss: 0.003062130883336067
step: 330, loss: 0.00022933761647436768
step: 340, loss: 0.0017091551562771201
step: 350, loss: 0.00019007985247299075
step: 360, loss: 0.00019421089382376522
step: 370, loss: 0.005149454344063997
step: 380, loss: 0.00013062912330497056
step: 390, loss: 0.00016104128735605627
step: 400, loss: 0.0011477865045890212
step: 410, loss: 0.025672471150755882
step: 420, loss: 0.2676667273044586
step: 430, loss: 0.022997641935944557
step: 440, loss: 0.0015437847469002008
step: 450, loss: 0.05885842815041542
step: 460, loss: 0.0021529337391257286
step: 470, loss: 0.01437652762979269
step: 480, loss: 0.010760004632174969
step: 490, loss: 0.002785362536087632
step: 500, loss: 0.011277731508016586
step: 510, loss: 0.0002524639421608299
step: 520, loss: 0.10837070643901825
step: 530, loss: 0.0029100358951836824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9554730983302412, f1=0.9533582089552238,
best_f1=0.9533582089552238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008805928751826286
step: 10, loss: 0.00015299160440918058
step: 20, loss: 9.999064059229568e-05
step: 30, loss: 0.00046462417230941355
step: 40, loss: 0.0005350231076590717
step: 50, loss: 0.12018483132123947
step: 60, loss: 0.005631075706332922
step: 70, loss: 0.01605963334441185
step: 80, loss: 0.0021060823928564787
step: 90, loss: 0.003661531023681164
step: 100, loss: 0.002358873374760151
step: 110, loss: 0.0008037617662921548
step: 120, loss: 0.0008631298551335931
step: 130, loss: 0.005398475099354982
step: 140, loss: 0.00042733788723126054
step: 150, loss: 8.914315549191087e-05
step: 160, loss: 0.07785189896821976
step: 170, loss: 0.0009069331572391093
step: 180, loss: 0.0007247595931403339
step: 190, loss: 0.021605994552373886
step: 200, loss: 0.0016427786322310567
step: 210, loss: 0.001185282482765615
step: 220, loss: 0.015528213232755661
step: 230, loss: 0.0018043743912130594
step: 240, loss: 0.012666157446801662
step: 250, loss: 0.08481988310813904
step: 260, loss: 0.022560030221939087
step: 270, loss: 0.11261910200119019
step: 280, loss: 0.001545955310575664
step: 290, loss: 0.004479157272726297
step: 300, loss: 0.002885487163439393
step: 310, loss: 0.06972266733646393
step: 320, loss: 0.00024367519654333591
step: 330, loss: 0.006698778364807367
step: 340, loss: 0.0010193770285695791
step: 350, loss: 0.0010262861615046859
step: 360, loss: 0.004072017036378384
step: 370, loss: 0.0013931307476013899
step: 380, loss: 0.0005715999868698418
step: 390, loss: 0.0023655088152736425
step: 400, loss: 0.010339499451220036
step: 410, loss: 0.000997216789983213
```

```

step: 420, loss: 0.00904251728206873
step: 430, loss: 0.0007492720033042133
step: 440, loss: 0.0013483710354194045
step: 450, loss: 0.1951691061258316
step: 460, loss: 0.004171645734459162
step: 470, loss: 0.004992588888853788
step: 480, loss: 0.006416228599846363
step: 490, loss: 0.012733405455946922
step: 500, loss: 0.014502678997814655
step: 510, loss: 0.051699597388505936
step: 520, loss: 0.00026561261620372534
step: 530, loss: 0.0001328872749581933
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9576271186440677, f1=0.9531396438612934,
best_f1=0.9531396438612934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029254595283418894
step: 10, loss: 0.0038688001222908497
step: 20, loss: 0.0004919188213534653
step: 30, loss: 0.016892682760953903
step: 40, loss: 0.007721523754298687
step: 50, loss: 0.0037023809272795916
step: 60, loss: 0.00637629721313715
step: 70, loss: 0.0004604989953804761
step: 80, loss: 0.0001277926639886573
step: 90, loss: 0.0011383788660168648
step: 100, loss: 0.000248657597694546
step: 110, loss: 8.234108099713922e-05
step: 120, loss: 9.878241689875722e-05
step: 130, loss: 7.6982170867268e-05
step: 140, loss: 0.0009830780327320099
step: 150, loss: 0.01052684336900711
step: 160, loss: 0.00016277135000564158
step: 170, loss: 0.04682914912700653
step: 180, loss: 0.0019337224075570703
step: 190, loss: 0.004491865634918213
step: 200, loss: 8.702500053914264e-05
step: 210, loss: 0.0006596195744350553
step: 220, loss: 0.00048420627717860043
step: 230, loss: 0.0001736154081299901

```



```

step: 240, loss: 0.08917026221752167
step: 250, loss: 0.009520660154521465
step: 260, loss: 0.005457150284200907
step: 270, loss: 0.020515285432338715
step: 280, loss: 0.003387751057744026
step: 290, loss: 0.00422905245795846
step: 300, loss: 8.239657472586259e-05
step: 310, loss: 0.0011233529075980186
step: 320, loss: 0.0052101160399615765
step: 330, loss: 0.001960531109943986
step: 340, loss: 0.012831032276153564
step: 350, loss: 0.0007497691549360752
step: 360, loss: 0.0007174399797804654
step: 370, loss: 0.02000519260764122
step: 380, loss: 0.08973944187164307
step: 390, loss: 0.0003242105303797871
step: 400, loss: 0.0019968394190073013
step: 410, loss: 0.0009551586117595434
step: 420, loss: 0.013531732372939587
step: 430, loss: 0.000689217122271657
step: 440, loss: 0.0004681370046455413
step: 450, loss: 0.0004196352092549205
step: 460, loss: 0.0003783678694162518
step: 470, loss: 0.09975484013557434
step: 480, loss: 0.004220213741064072
step: 490, loss: 0.00012764618440996855
step: 500, loss: 0.0021471644286066294
step: 510, loss: 0.00011181776790181175
step: 520, loss: 6.395364471245557e-05
step: 530, loss: 0.0020034443587064743
/usr/local/lib/python3.7/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=0.9573283858998145, f1=0.9572093023255814,
best_f1=0.9531396438612934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018713832832872868
step: 10, loss: 0.0035141294356435537
step: 20, loss: 0.006907516624778509
step: 30, loss: 0.00021104906045366079
step: 40, loss: 2.6534442440606654e-05
step: 50, loss: 7.880826888140291e-05
step: 60, loss: 0.00019013250130228698
step: 70, loss: 7.542569801444188e-05
step: 80, loss: 0.08561669290065765
step: 90, loss: 9.386649617226794e-05

```

```
step: 100, loss: 0.00013978594506625086
step: 110, loss: 0.0003320930409245193
step: 120, loss: 0.0006007052143104374
step: 130, loss: 0.0624796487390995
step: 140, loss: 0.00023461252567358315
step: 150, loss: 0.00015787710435688496
step: 160, loss: 0.0002717781171668321
step: 170, loss: 0.0024979598820209503
step: 180, loss: 0.0026094766799360514
step: 190, loss: 0.022256391122937202
step: 200, loss: 0.004626184236258268
step: 210, loss: 0.052879393100738525
step: 220, loss: 0.00030278362100943923
step: 230, loss: 0.057097017765045166
step: 240, loss: 0.002733167726546526
step: 250, loss: 0.008104421198368073
step: 260, loss: 0.00014398767962120473
step: 270, loss: 0.01756487600505352
step: 280, loss: 0.0002082669670926407
step: 290, loss: 0.0038105216808617115
step: 300, loss: 6.030082658980973e-05
step: 310, loss: 0.0004104285908397287
step: 320, loss: 0.0005687664961442351
step: 330, loss: 0.0005053599597886205
step: 340, loss: 0.015610876493155956
step: 350, loss: 0.00029707487556152046
step: 360, loss: 0.0005045363795943558
step: 370, loss: 0.062024228274822235
step: 380, loss: 9.771347686182708e-05
step: 390, loss: 0.03356403484940529
step: 400, loss: 0.003998153377324343
step: 410, loss: 0.005388285499066114
step: 420, loss: 6.0722701164195314e-05
step: 430, loss: 0.0033702438231557608
step: 440, loss: 0.002556754508987069
step: 450, loss: 0.0018321149982511997
step: 460, loss: 0.0010578003711998463
step: 470, loss: 0.052806369960308075
step: 480, loss: 0.003868949366733432
step: 490, loss: 0.02206565998494625
step: 500, loss: 0.0008394923061132431
step: 510, loss: 0.001421273103915155
step: 520, loss: 0.0015835545491427183
step: 530, loss: 0.017970236018300056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.96040987424313, f1=0.9563197026022305,

best_f1=0.9563197026022305

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

An input tensor was not cuda.

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

step: 0, loss: 4.605873618857004e-05

step: 10, loss: 0.0014521260745823383

step: 20, loss: 4.9546499212738127e-05

step: 30, loss: 0.07641471922397614

step: 40, loss: 0.000257775274803862

step: 50, loss: 0.00019196863286197186

step: 60, loss: 0.0013242432614788413

step: 70, loss: 0.023839568719267845

step: 80, loss: 0.004247293807566166

step: 90, loss: 0.04140244424343109

step: 100, loss: 0.00025953122531063855

step: 110, loss: 0.0771859660744667

step: 120, loss: 0.0008400884689763188

step: 130, loss: 0.0018931300146505237

step: 140, loss: 0.002497555920854211

step: 150, loss: 0.0004980582161806524

step: 160, loss: 0.00012568912643473595

step: 170, loss: 0.0016546191181987524

step: 180, loss: 0.04211907833814621

step: 190, loss: 0.0005490185576491058

step: 200, loss: 0.0009937919676303864

step: 210, loss: 0.0006995244766585529

step: 220, loss: 0.00019068081746809185

step: 230, loss: 0.0003451380762271583

step: 240, loss: 9.298104123445228e-05

step: 250, loss: 0.0003904499753843993

step: 260, loss: 0.0008768292027525604

step: 270, loss: 0.0018940020818263292

step: 280, loss: 0.02698264829814434

step: 290, loss: 0.0001036910543916747

step: 300, loss: 0.0004250903148204088

step: 310, loss: 0.004662413150072098

step: 320, loss: 4.4441774662118405e-05

step: 330, loss: 0.00044722514576278627

step: 340, loss: 0.0007471739081665874

step: 350, loss: 0.02916925959289074

step: 360, loss: 3.4605942346388474e-05

step: 370, loss: 7.833060226403177e-05

step: 380, loss: 0.00023208072525449097

step: 390, loss: 1.878994589787908e-05

```

step: 400, loss: 3.1641520763514563e-05
step: 410, loss: 7.474374433513731e-05
step: 420, loss: 6.778911483706906e-05
step: 430, loss: 0.010518679395318031
step: 440, loss: 6.98822841513902e-05
step: 450, loss: 0.00788886845111847
step: 460, loss: 0.00022668749443255365
step: 470, loss: 0.0007901416392996907
step: 480, loss: 0.00013913463044445962
step: 490, loss: 0.0317717120051384
step: 500, loss: 0.0019368698121979833
step: 510, loss: 0.003863454330712557
step: 520, loss: 0.006556525826454163
step: 530, loss: 0.004008985590189695
/usr/local/lib/python3.7/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=0.9554317548746518, f1=0.9537892791127542,
best_f1=0.9563197026022305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038177547976374626
step: 10, loss: 0.0008736880263313651
step: 20, loss: 0.0011414886685088277
step: 30, loss: 0.017534738406538963
step: 40, loss: 0.0001980713423108682
step: 50, loss: 0.00140077481046319
step: 60, loss: 0.00679645175114274
step: 70, loss: 2.1874426238355227e-05
step: 80, loss: 0.000111659690446686
step: 90, loss: 4.85408709209878e-05
step: 100, loss: 0.003102667164057493
step: 110, loss: 0.030131744220852852
step: 120, loss: 6.936270801816136e-05
step: 130, loss: 3.6143435863777995e-05
step: 140, loss: 8.214667468564585e-05
step: 150, loss: 3.381639544386417e-05
step: 160, loss: 0.053715743124485016
step: 170, loss: 4.46524572907947e-05
step: 180, loss: 6.139290053397417e-05
step: 190, loss: 1.2133091331634205e-05
step: 200, loss: 0.08763670921325684
step: 210, loss: 0.0020450300071388483
step: 220, loss: 0.001702994224615395
step: 230, loss: 0.006747964769601822
step: 240, loss: 4.849625111091882e-05
step: 250, loss: 0.0072012548334896564

```

```

step: 260, loss: 0.0019855028949677944
step: 270, loss: 9.873494127532467e-05
step: 280, loss: 0.024686094373464584
step: 290, loss: 0.002708428306505084
step: 300, loss: 0.0017962995916604996
step: 310, loss: 0.07188237458467484
step: 320, loss: 0.01104320865124464
step: 330, loss: 0.0004876343300566077
step: 340, loss: 4.3566393287619576e-05
step: 350, loss: 0.023602521046996117
step: 360, loss: 2.5386540073668584e-05
step: 370, loss: 0.00266038766130805
step: 380, loss: 0.001631335006095469
step: 390, loss: 2.4481947548338212e-05
step: 400, loss: 0.0006924844346940517
step: 410, loss: 5.4786025430075824e-05
step: 420, loss: 2.6798397811944596e-05
step: 430, loss: 0.0012680967338383198
step: 440, loss: 3.4250948374392465e-05
step: 450, loss: 8.886036084732041e-05
step: 460, loss: 3.5575903893914074e-05
step: 470, loss: 0.0036624560598284006
step: 480, loss: 3.469291186775081e-05
step: 490, loss: 0.0004927460104227066
step: 500, loss: 0.0001596063666511327
step: 510, loss: 1.9065275409957394e-05
step: 520, loss: 0.0012286544078961015
step: 530, loss: 0.0014592010993510485
/usr/local/lib/python3.7/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=0.9543147208121827, f1=0.9504132231404958,
best_f1=0.9563197026022305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.821250407374464e-05
step: 10, loss: 2.669029163371306e-05
step: 20, loss: 1.4673624718852807e-05
step: 30, loss: 0.00014434271724894643
step: 40, loss: 2.0119547116337344e-05
step: 50, loss: 3.698427099152468e-05
step: 60, loss: 2.473663153068628e-05
step: 70, loss: 2.6746503863250837e-05
step: 80, loss: 0.000629142508842051
step: 90, loss: 0.0011433395557105541
step: 100, loss: 3.07519621856045e-05
step: 110, loss: 2.216482607764192e-05

```

```

step: 120, loss: 4.417432501213625e-05
step: 130, loss: 1.8432056094752625e-05
step: 140, loss: 2.0112058336962946e-05
step: 150, loss: 2.0030041923746467e-05
step: 160, loss: 4.187755621387623e-05
step: 170, loss: 6.774705980205908e-05
step: 180, loss: 2.9063965484965593e-05
step: 190, loss: 0.00023558875545859337
step: 200, loss: 1.3779523214907385e-05
step: 210, loss: 1.4762805221835151e-05
step: 220, loss: 0.015073812566697598
step: 230, loss: 6.589958957192721e-06
step: 240, loss: 0.00279112602584064
step: 250, loss: 5.6668119214009494e-05
step: 260, loss: 0.010234616696834564
step: 270, loss: 0.004221462644636631
step: 280, loss: 0.00015197416360024363
step: 290, loss: 0.00014057991211302578
step: 300, loss: 7.640579860890284e-05
step: 310, loss: 0.07907895743846893
step: 320, loss: 5.325349047780037e-05
step: 330, loss: 1.2572600098792464e-05
step: 340, loss: 0.00031006347853690386
step: 350, loss: 0.0002726311795413494
step: 360, loss: 0.0001606822042958811
step: 370, loss: 0.004674531985074282
step: 380, loss: 9.974322892958298e-05
step: 390, loss: 0.0020829932764172554
step: 400, loss: 0.0034079623874276876
step: 410, loss: 0.0010990883456543088
step: 420, loss: 1.5355135474237613e-05
step: 430, loss: 0.002371913054957986
step: 440, loss: 3.373607978574e-05
step: 450, loss: 0.0011262421030551195
step: 460, loss: 0.0003257741336710751
step: 470, loss: 0.0005007191793993115
step: 480, loss: 3.187083711964078e-05
step: 490, loss: 2.6708330551628023e-05
step: 500, loss: 0.001286305021494627
step: 510, loss: 2.5475519578321837e-05
step: 520, loss: 8.303527283715084e-05
step: 530, loss: 0.029031138867139816
/usr/local/lib/python3.7/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=0.9554730983302412, f1=0.9562816382880809,
best_f1=0.9563197026022305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0036056467797607183
step: 10, loss: 0.002344787120819092
step: 20, loss: 0.006605622358620167
step: 30, loss: 0.002812011865898967
step: 40, loss: 0.0005254459101706743
step: 50, loss: 0.009732276201248169
step: 60, loss: 3.2426047255285084e-05
step: 70, loss: 0.00047697001718916
step: 80, loss: 0.0035042655654251575
step: 90, loss: 0.0008887287694960833
step: 100, loss: 0.0016560029471293092
step: 110, loss: 1.7206395568791777e-05
step: 120, loss: 4.6329623728524894e-05
step: 130, loss: 0.0014536798698827624
step: 140, loss: 0.0002357687772018835
step: 150, loss: 2.6542276827967726e-05
step: 160, loss: 0.0008285453077405691
step: 170, loss: 2.2257088858168572e-05
step: 180, loss: 1.9523395167198032e-05
step: 190, loss: 1.760534905770328e-05
step: 200, loss: 0.0026933352928608656
step: 210, loss: 1.3235713595349807e-05
step: 220, loss: 2.1195237422944047e-05
step: 230, loss: 1.2069644981238525e-05
step: 240, loss: 0.0008882149122655392
step: 250, loss: 1.2836902897106484e-05
step: 260, loss: 8.877238542481791e-06
step: 270, loss: 0.0153993284329772
step: 280, loss: 1.8435990568832494e-05
step: 290, loss: 0.00042425774154253304
step: 300, loss: 0.0021808664314448833
step: 310, loss: 2.2476850062957965e-05
step: 320, loss: 6.336501974146813e-05
step: 330, loss: 0.1318863332271576
step: 340, loss: 2.376526390435174e-05
step: 350, loss: 0.0004908909322693944
step: 360, loss: 2.3170496206148528e-05
step: 370, loss: 0.0004349159717094153
step: 380, loss: 0.0009984917705878615
step: 390, loss: 6.962419138289988e-05
step: 400, loss: 1.2743958905048203e-05
step: 410, loss: 2.7942143788095564e-05
step: 420, loss: 4.5493958168663085e-05
step: 430, loss: 5.212083488004282e-05
step: 440, loss: 0.001296731410548091
step: 450, loss: 0.0005933645297773182
```

```

step: 460, loss: 5.007824438507669e-05
step: 470, loss: 0.0011918519157916307
step: 480, loss: 0.0114121213555336
step: 490, loss: 3.881171505781822e-05
step: 500, loss: 4.122056270716712e-05
step: 510, loss: 0.013839700259268284
step: 520, loss: 0.0009889379143714905
step: 530, loss: 3.645435572252609e-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=0.9577464788732395, f1=0.9570093457943926,
best_f1=0.9563197026022305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.0611576474038884e-05
step: 10, loss: 2.764091368590016e-05
step: 20, loss: 5.974624946247786e-05
step: 30, loss: 4.3821353756356984e-05
step: 40, loss: 0.0018662161892279983
step: 50, loss: 0.00809817761182785
step: 60, loss: 0.0034376431722193956
step: 70, loss: 0.00031291780760511756
step: 80, loss: 0.00011580019781831652
step: 90, loss: 2.593091267044656e-05
step: 100, loss: 2.043984568445012e-05
step: 110, loss: 0.00021241497597657144
step: 120, loss: 2.544628114264924e-05
step: 130, loss: 2.2056519810575992e-05
step: 140, loss: 0.0009605559753254056
step: 150, loss: 0.00200461084023118
step: 160, loss: 0.0007003420032560825
step: 170, loss: 0.0022620712406933308
step: 180, loss: 0.00041472777957096696
step: 190, loss: 0.002209392609074712
step: 200, loss: 1.6044496078393422e-05
step: 210, loss: 1.9452496417216025e-05
step: 220, loss: 1.3999425391375553e-05
step: 230, loss: 0.0013015283038839698
step: 240, loss: 4.529683792497963e-05
step: 250, loss: 0.00032064903643913567
step: 260, loss: 1.809333116398193e-05
step: 270, loss: 0.0001559526426717639
step: 280, loss: 2.8067457606084645e-05
step: 290, loss: 1.3313811905391049e-05
step: 300, loss: 1.634255750104785e-05
step: 310, loss: 4.988313594367355e-05

```



```

step: 320, loss: 1.5653333321097307e-05
step: 330, loss: 1.9378656361368485e-05
step: 340, loss: 0.002458173083141446
step: 350, loss: 7.670319973840378e-06
step: 360, loss: 0.09766218811273575
step: 370, loss: 0.014275174587965012
step: 380, loss: 0.0005486871814355254
step: 390, loss: 1.119064927479485e-05
step: 400, loss: 2.080139529425651e-05
step: 410, loss: 2.4827902961988002e-05
step: 420, loss: 1.5347912267316133e-05
step: 430, loss: 4.5060933189233765e-05
step: 440, loss: 9.752708137966692e-06
step: 450, loss: 1.7363152437610552e-05
step: 460, loss: 0.002323816530406475
step: 470, loss: 0.0014473971677944064
step: 480, loss: 7.92363516666228e-06
step: 490, loss: 8.339570194948465e-05
step: 500, loss: 9.17530087463092e-06
step: 510, loss: 1.0847922567336354e-05
step: 520, loss: 1.7623846360947937e-05
step: 530, loss: 1.577602233737707e-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=0.9572970436414829, f1=0.9563197026022305,
best_f1=0.9563197026022305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00043665169505402446
step: 10, loss: 0.0014907210133969784
step: 20, loss: 8.903367415769026e-06
step: 30, loss: 0.0018165666842833161
step: 40, loss: 0.0007004959043115377
step: 50, loss: 0.0017837748164311051
step: 60, loss: 1.3395953828876372e-05
step: 70, loss: 2.688166387088131e-05
step: 80, loss: 1.6200810932787135e-05
step: 90, loss: 9.767611118149944e-06
step: 100, loss: 1.0382245818618685e-05
step: 110, loss: 0.0017144330777227879
step: 120, loss: 1.3459228284773417e-05
step: 130, loss: 0.00016623873671051115
step: 140, loss: 1.725521360640414e-05
step: 150, loss: 0.0025161050725728273
step: 160, loss: 1.9482773495838046e-05
step: 170, loss: 0.008569402620196342

```

```

step: 180, loss: 8.844250260153785e-05
step: 190, loss: 0.00010672890493879095
step: 200, loss: 4.8394271288998425e-05
step: 210, loss: 0.0008533119107596576
step: 220, loss: 4.0578015614300966e-05
step: 230, loss: 5.895836147828959e-05
step: 240, loss: 0.0012159014586359262
step: 250, loss: 0.0013803828042000532
step: 260, loss: 6.350800686050206e-05
step: 270, loss: 0.0017233012476935983
step: 280, loss: 5.526846871362068e-05
step: 290, loss: 2.465342367941048e-05
step: 300, loss: 2.1736585040343925e-05
step: 310, loss: 0.000345254666171968
step: 320, loss: 2.447428232699167e-05
step: 330, loss: 0.0006622858927585185
step: 340, loss: 0.000693367444910109
step: 350, loss: 1.7288626622757874e-05
step: 360, loss: 2.3137248717830516e-05
step: 370, loss: 0.004272539634257555
step: 380, loss: 0.0003071839455515146
step: 390, loss: 2.692903217393905e-05
step: 400, loss: 0.001260262099094689
step: 410, loss: 2.271908124384936e-05
step: 420, loss: 2.5736753741512075e-05
step: 430, loss: 0.002895694226026535
step: 440, loss: 3.88023690902628e-05
step: 450, loss: 1.6182420949917287e-05
step: 460, loss: 0.01956780254840851
step: 470, loss: 0.002916189609095454
step: 480, loss: 1.3794532605970744e-05
step: 490, loss: 1.4673527402919717e-05
step: 500, loss: 0.0018499715952202678
step: 510, loss: 0.023327697068452835
step: 520, loss: 0.000691303750500083
step: 530, loss: 0.0004809439124073833
/usr/local/lib/python3.7/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=0.9552520018841262, f1=0.9536299765807962,
best_f1=0.9563197026022305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.859968324424699e-05
step: 10, loss: 0.002064033877104521
step: 20, loss: 0.0018096122657880187
step: 30, loss: 0.00978065188974142

```

step: 40, loss: 1.6502650396432728e-05
step: 50, loss: 0.0011947317980229855
step: 60, loss: 3.300513344584033e-05
step: 70, loss: 2.793934618239291e-05
step: 80, loss: 1.9214554413338192e-05
step: 90, loss: 1.9016586520592682e-05
step: 100, loss: 2.3243421310326084e-05
step: 110, loss: 0.0009299884550273418
step: 120, loss: 1.3977039088786114e-05
step: 130, loss: 2.5057488528545946e-05
step: 140, loss: 4.3414071114966646e-05
step: 150, loss: 1.981732748390641e-05
step: 160, loss: 1.849159707489889e-05
step: 170, loss: 2.3567918105982244e-05
step: 180, loss: 0.000285326357698068
step: 190, loss: 0.001080971211194992
step: 200, loss: 0.001184558030217886
step: 210, loss: 2.3606027752975933e-05
step: 220, loss: 3.2723335607443005e-05
step: 230, loss: 0.00011359874770278111
step: 240, loss: 1.654006882745307e-05
step: 250, loss: 3.037398710148409e-05
step: 260, loss: 1.5683230230933987e-05
step: 270, loss: 1.4249077139538713e-05
step: 280, loss: 1.4796662981098052e-05
step: 290, loss: 0.0006021157023496926
step: 300, loss: 2.9111602998455055e-05
step: 310, loss: 0.035305656492710114
step: 320, loss: 2.7885207600775175e-05
step: 330, loss: 2.4794722776277922e-05
step: 340, loss: 2.814975414366927e-05
step: 350, loss: 0.002547457115724683
step: 360, loss: 8.826491830404848e-05
step: 370, loss: 1.442044049326796e-05
step: 380, loss: 2.328165282960981e-05
step: 390, loss: 5.627125210594386e-05
step: 400, loss: 4.530617661657743e-05
step: 410, loss: 5.113680163049139e-05
step: 420, loss: 3.012420711456798e-05
step: 430, loss: 1.0900093911914155e-05
step: 440, loss: 3.3390639146091416e-05
step: 450, loss: 0.009815510362386703
step: 460, loss: 0.0018004968296736479
step: 470, loss: 1.1231663847866002e-05
step: 480, loss: 0.0025207828730344772
step: 490, loss: 3.394148370716721e-05
step: 500, loss: 0.0031054662540555
step: 510, loss: 1.2114534911233932e-05

```

step: 520, loss: 2.9795410227961838e-05
step: 530, loss: 1.2710558621620294e-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=0.9559512652296158, f1=0.9558755225267069,
best_f1=0.9563197026022305
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, 168.48it/s]
load_f1 = 0.9573283858998145
real_f1 = 0.957169459962756
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, 139.83it/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: 0.4997784495353699
step: 10, loss: 0.4656345248222351
step: 20, loss: 0.47685062885284424
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.34332355856895447
step: 40, loss: 0.335739403963089
step: 50, loss: 0.38646814227104187
step: 60, loss: 0.44769996404647827
step: 70, loss: 0.3120538592338562
step: 80, loss: 0.32581013441085815
step: 90, loss: 0.29789474606513977
step: 100, loss: 0.24808458983898163
step: 110, loss: 0.26842576265335083
step: 120, loss: 0.3988053500652313
step: 130, loss: 0.3041296601295471
step: 140, loss: 0.4876771569252014
step: 150, loss: 0.3902084231376648
step: 160, loss: 0.4914036989212036
step: 170, loss: 0.22991308569908142
step: 180, loss: 0.37427037954330444
step: 190, loss: 0.5858056545257568
step: 200, loss: 0.2500547766685486
step: 210, loss: 0.40996456146240234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4160206718346253, f1=0.41138421733505826,
best_f1=0.41138421733505826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3103681802749634
step: 10, loss: 0.1728707253932953
step: 20, loss: 0.5223801136016846
step: 30, loss: 0.4880716800689697
step: 40, loss: 0.5327901840209961
step: 50, loss: 0.1712644100189209
step: 60, loss: 0.338190495967865
step: 70, loss: 0.2709077000617981
step: 80, loss: 0.22010907530784607
step: 90, loss: 0.21848569810390472
step: 100, loss: 0.4921967387199402
step: 110, loss: 0.2584613263607025
step: 120, loss: 0.16934502124786377
step: 130, loss: 0.16374556720256805
step: 140, loss: 0.1942596435546875
step: 150, loss: 0.29821690917015076
step: 160, loss: 0.13708557188510895
step: 170, loss: 0.3272401690483093
step: 180, loss: 0.16796453297138214
step: 190, loss: 0.40067583322525024
step: 200, loss: 0.0807720348238945
step: 210, loss: 0.20464879274368286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6457564575645757, f1=0.6983546617915904,
best_f1=0.6983546617915904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0332111231982708
step: 10, loss: 0.06384794414043427
step: 20, loss: 0.24301281571388245
step: 30, loss: 0.05699887499213219
step: 40, loss: 0.3404531478881836

```

```

step: 50, loss: 0.24990297853946686
step: 60, loss: 0.2334730625152588
step: 70, loss: 0.10228259116411209
step: 80, loss: 0.16271483898162842
step: 90, loss: 0.04140235111117363
step: 100, loss: 0.1693168729543686
step: 110, loss: 0.1084320917725563
step: 120, loss: 0.14523983001708984
step: 130, loss: 0.22313228249549866
step: 140, loss: 0.11181306838989258
step: 150, loss: 0.15970538556575775
step: 160, loss: 0.19091610610485077
step: 170, loss: 0.2800447940826416
step: 180, loss: 0.24112765491008759
step: 190, loss: 0.059081610292196274
step: 200, loss: 0.13141189515590668
step: 210, loss: 0.19949685037136078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6911196911196912, f1=0.7082494969818913,
best_f1=0.7082494969818913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04453461989760399
step: 10, loss: 0.11181666702032089
step: 20, loss: 0.18453465402126312
step: 30, loss: 0.06568850576877594
step: 40, loss: 0.05739782005548477
step: 50, loss: 0.14027006924152374
step: 60, loss: 0.1767881214618683
step: 70, loss: 0.0758463516831398
step: 80, loss: 0.22720059752464294
step: 90, loss: 0.1435006558895111
step: 100, loss: 0.3244342505931854
step: 110, loss: 0.5415915250778198
step: 120, loss: 0.18957234919071198
step: 130, loss: 0.35983774065971375
step: 140, loss: 0.16530925035476685
step: 150, loss: 0.056995194405317307
step: 160, loss: 0.11315537989139557
step: 170, loss: 0.2191479653120041
step: 180, loss: 0.05621800199151039

```



```

step: 190, loss: 0.12173139303922653
step: 200, loss: 0.08879265189170837
step: 210, loss: 0.3002066910266876
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.710204081632653, f1=0.7310924369747899,
best_f1=0.7310924369747899
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15701664984226227
step: 10, loss: 0.07667861133813858
step: 20, loss: 0.11715889722108841
step: 30, loss: 0.04532928764820099
step: 40, loss: 0.0779867097735405
step: 50, loss: 0.1558697372674942
step: 60, loss: 0.1701226383447647
step: 70, loss: 0.17431102693080902
step: 80, loss: 0.065669484436512
step: 90, loss: 0.08612523972988129
step: 100, loss: 0.033806212246418
step: 110, loss: 0.1270151436328888
step: 120, loss: 0.10064753144979477
step: 130, loss: 0.1573137789964676
step: 140, loss: 0.12412891536951065
step: 150, loss: 0.2166052907705307
step: 160, loss: 0.05777968466281891
step: 170, loss: 0.031967390328645706
step: 180, loss: 0.050439994782209396
step: 190, loss: 0.15183013677597046
step: 200, loss: 0.05180192366242409
step: 210, loss: 0.0989755317568779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7122736418511065, f1=0.728744939271255,
best_f1=0.728744939271255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045692481100559235
step: 10, loss: 0.16949918866157532
step: 20, loss: 0.09874293953180313
step: 30, loss: 0.021769532933831215
step: 40, loss: 0.028911059722304344
step: 50, loss: 0.06879817694425583
step: 60, loss: 0.13590644299983978
step: 70, loss: 0.13062705099582672
step: 80, loss: 0.11682287603616714
step: 90, loss: 0.09902249276638031
step: 100, loss: 0.09262631833553314
step: 110, loss: 0.07855043560266495
step: 120, loss: 0.031691163778305054
step: 130, loss: 0.017338531091809273
step: 140, loss: 0.09319048374891281
step: 150, loss: 0.11053935438394547
step: 160, loss: 0.05598857253789902
step: 170, loss: 0.06845976412296295
step: 180, loss: 0.19418445229530334
step: 190, loss: 0.04210837185382843
step: 200, loss: 0.10937267541885376
step: 210, loss: 0.09174823760986328
/usr/local/lib/python3.7/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=0.7061143984220907, f1=0.7250996015936255,
best_f1=0.728744939271255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04758010804653168
step: 10, loss: 0.018606150522828102
step: 20, loss: 0.00905058067291975
step: 30, loss: 0.02004903368651867
step: 40, loss: 0.013970812782645226
step: 50, loss: 0.08264511078596115
step: 60, loss: 0.05266164243221283
step: 70, loss: 0.041080787777900696
step: 80, loss: 0.029046911746263504
step: 90, loss: 0.04986025020480156
step: 100, loss: 0.05085810273885727
step: 110, loss: 0.25183987617492676
step: 120, loss: 0.03088214434683323
step: 130, loss: 0.09294085949659348
step: 140, loss: 0.04300002381205559
step: 150, loss: 0.06541037559509277
step: 160, loss: 0.30397942662239075

```

```

step: 170, loss: 0.12410939484834671
step: 180, loss: 0.04396304115653038
step: 190, loss: 0.10463211685419083
step: 200, loss: 0.02998918481171131
step: 210, loss: 0.2101721465587616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7309236947791166, f1=0.7117296222664016,
best_f1=0.7117296222664016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22837881743907928
step: 10, loss: 0.03381447494029999
step: 20, loss: 0.07352638244628906
step: 30, loss: 0.11989682912826538
step: 40, loss: 0.03709065169095993
step: 50, loss: 0.014508034102618694
step: 60, loss: 0.06315556168556213
step: 70, loss: 0.15307800471782684
step: 80, loss: 0.09737871587276459
step: 90, loss: 0.09127403050661087
step: 100, loss: 0.1514074206352234
step: 110, loss: 0.01852744072675705
step: 120, loss: 0.018212901428341866
step: 130, loss: 0.017608972266316414
step: 140, loss: 0.07393515110015869
step: 150, loss: 0.2316170334815979
step: 160, loss: 0.2337973266839981
step: 170, loss: 0.2795726954936981
step: 180, loss: 0.09905938059091568
step: 190, loss: 0.018689891323447227
step: 200, loss: 0.059892866760492325
step: 210, loss: 0.06462769955396652
/usr/local/lib/python3.7/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=0.7050092764378478, f1=0.7134935304990756,
best_f1=0.7117296222664016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059221312403678894

```

```

step: 10, loss: 0.1036941260099411
step: 20, loss: 0.09761442244052887
step: 30, loss: 0.02563839592039585
step: 40, loss: 0.01168979424983263
step: 50, loss: 0.05203579366207123
step: 60, loss: 0.007423983421176672
step: 70, loss: 0.07336238771677017
step: 80, loss: 0.05737227946519852
step: 90, loss: 0.05514885485172272
step: 100, loss: 0.057798948138952255
step: 110, loss: 0.11102315783500671
step: 120, loss: 0.13505995273590088
step: 130, loss: 0.02185116708278656
step: 140, loss: 0.15832284092903137
step: 150, loss: 0.007011116482317448
step: 160, loss: 0.13890954852104187
step: 170, loss: 0.16862048208713531
step: 180, loss: 0.050466541200876236
step: 190, loss: 0.10284482687711716
step: 200, loss: 0.02838117629289627
step: 210, loss: 0.015515903010964394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.7345309381237525, f1=0.716297786720322,
best_f1=0.716297786720322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01644711196422577
step: 10, loss: 0.04396677017211914
step: 20, loss: 0.011871062219142914
step: 30, loss: 0.005574194714426994
step: 40, loss: 0.03911060094833374
step: 50, loss: 0.026123039424419403
step: 60, loss: 0.01834218017756939
step: 70, loss: 0.09295542538166046
step: 80, loss: 0.09258366376161575
step: 90, loss: 0.03710775077342987
step: 100, loss: 0.05954045057296753
step: 110, loss: 0.02907741256058216
step: 120, loss: 0.030494289472699165
step: 130, loss: 0.01577160321176052
step: 140, loss: 0.025567835196852684

```

```

step: 150, loss: 0.03880181163549423
step: 160, loss: 0.020688993856310844
step: 170, loss: 0.05884844437241554
step: 180, loss: 0.02674846164882183
step: 190, loss: 0.018926328048110008
step: 200, loss: 0.13312798738479614
step: 210, loss: 0.0268105436116457
/usr/local/lib/python3.7/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=0.7258064516129032, f1=0.7063655030800821,
best_f1=0.716297786720322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08044479042291641
step: 10, loss: 0.005547824781388044
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.012198545970022678
step: 30, loss: 0.013573585078120232
step: 40, loss: 0.036665432155132294
step: 50, loss: 0.01550429966300726
step: 60, loss: 0.04154594615101814
step: 70, loss: 0.08609115332365036
step: 80, loss: 0.20327863097190857
step: 90, loss: 0.18224632740020752
step: 100, loss: 0.05654916167259216
step: 110, loss: 0.006468122825026512
step: 120, loss: 0.12317436188459396
step: 130, loss: 0.015762845054268837
step: 140, loss: 0.05158815160393715
step: 150, loss: 0.04572586342692375
step: 160, loss: 0.0052749100141227245
step: 170, loss: 0.02937978506088257
step: 180, loss: 0.023280257359147072
step: 190, loss: 0.04587910324335098
step: 200, loss: 0.0057036359794437885
step: 210, loss: 0.15006250143051147
/usr/local/lib/python3.7/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=0.7203389830508475, f1=0.7161016949152543,
best_f1=0.716297786720322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007312262896448374
step: 10, loss: 0.02764139324426651

```

```

step: 20, loss: 0.0965629443526268
step: 30, loss: 0.01343566458672285
step: 40, loss: 0.0006095269345678389
step: 50, loss: 0.002682638820260763
step: 60, loss: 0.0010271945502609015
step: 70, loss: 0.07629137486219406
step: 80, loss: 0.08200852572917938
step: 90, loss: 0.02623436599969864
step: 100, loss: 0.010121689178049564
step: 110, loss: 0.055077798664569855
step: 120, loss: 0.0032475674524903297
step: 130, loss: 0.15984076261520386
step: 140, loss: 0.03373817354440689
step: 150, loss: 0.003351217834278941
step: 160, loss: 0.057465359568595886
step: 170, loss: 0.03353790566325188
step: 180, loss: 0.02303279936313629
step: 190, loss: 0.0071339975111186504
step: 200, loss: 0.009317029267549515
step: 210, loss: 0.009510310366749763
/usr/local/lib/python3.7/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=0.7218045112781954, f1=0.7327102803738318,
best_f1=0.716297786720322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018156785517930984
step: 10, loss: 0.01649269461631775
step: 20, loss: 0.012174537405371666
step: 30, loss: 0.008542865514755249
step: 40, loss: 0.011760726571083069
step: 50, loss: 0.17954924702644348
step: 60, loss: 0.015784980729222298
step: 70, loss: 0.015566874295473099
step: 80, loss: 0.18662667274475098
step: 90, loss: 0.011044020764529705
step: 100, loss: 0.015752039849758148
step: 110, loss: 0.11691247671842575
step: 120, loss: 0.1963028907775879
step: 130, loss: 0.0051061613485217094
step: 140, loss: 0.03051452711224556
step: 150, loss: 0.012442795559763908
step: 160, loss: 0.061521075665950775
step: 170, loss: 0.008800542913377285
step: 180, loss: 0.008415895514190197
step: 190, loss: 0.016417773440480232

```

```

step: 200, loss: 0.02922092378139496
step: 210, loss: 0.05432334914803505
/usr/local/lib/python3.7/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=0.719851576994434, f1=0.7129455909943716,
best_f1=0.716297786720322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011566640809178352
step: 10, loss: 0.05030880868434906
step: 20, loss: 0.13785246014595032
step: 30, loss: 0.015678219497203827
step: 40, loss: 0.04271557554602623
step: 50, loss: 0.009783503599464893
step: 60, loss: 0.07861576974391937
step: 70, loss: 0.007154738996177912
step: 80, loss: 0.07702022790908813
step: 90, loss: 0.014592926949262619
step: 100, loss: 0.011969235725700855
step: 110, loss: 0.07906045019626617
step: 120, loss: 0.002665066858753562
step: 130, loss: 0.01525910571217537
step: 140, loss: 0.01165236346423626
step: 150, loss: 0.06072255223989487
step: 160, loss: 0.03112981468439102
step: 170, loss: 0.09228120744228363
step: 180, loss: 0.0034336168318986893
step: 190, loss: 0.06297925114631653
step: 200, loss: 0.0072898827493190765
step: 210, loss: 0.002298880135640502
/usr/local/lib/python3.7/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=0.732824427480916, f1=0.7101727447216889,
best_f1=0.716297786720322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025754185393452644
step: 10, loss: 0.003204546170309186
step: 20, loss: 0.04509308561682701
step: 30, loss: 0.013174665160477161
step: 40, loss: 0.02271537482738495
step: 50, loss: 0.2003769874572754
step: 60, loss: 0.0389050655066967
step: 70, loss: 0.005421274341642857

```

```

step: 80, loss: 0.01394482608884573
step: 90, loss: 0.005056883208453655
step: 100, loss: 0.00412800582125783
step: 110, loss: 0.11073713004589081
step: 120, loss: 0.040080081671476364
step: 130, loss: 0.01660160906612873
step: 140, loss: 0.021390236914157867
step: 150, loss: 0.020409682765603065
step: 160, loss: 0.08972062915563583
step: 170, loss: 0.049627356231212616
step: 180, loss: 0.011117284186184406
step: 190, loss: 0.02199767529964447
step: 200, loss: 0.004120455589145422
step: 210, loss: 0.07783617824316025
/usr/local/lib/python3.7/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=0.725897920604915, f1=0.7129455909943716,
best_f1=0.716297786720322
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 [00:09, 237.94it/s]
load_f1 = 0.72936660268714
real_f1 = 0.7279693486590039
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, 138.68it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.47408121824264526
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41543900966644287
step: 20, loss: 0.2925112843513489
step: 30, loss: 0.3397584855556488
step: 40, loss: 0.26356980204582214
step: 50, loss: 0.30494001507759094
step: 60, loss: 0.5150033235549927
step: 70, loss: 0.4702031910419464
step: 80, loss: 0.2057049721479416
step: 90, loss: 0.3231757879257202
step: 100, loss: 0.4296358823776245
step: 110, loss: 0.22888246178627014
step: 120, loss: 0.3001486659049988
step: 130, loss: 0.36732906103134155
step: 140, loss: 0.15926755964756012

```

```

step: 150, loss: 0.27805620431900024
step: 160, loss: 0.19451850652694702
step: 170, loss: 0.4545954167842865
step: 180, loss: 0.1180264875292778
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.14102450013160706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6363636363636364, f1=0.6463104325699746,
best_f1=0.6463104325699746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.43399086594581604
step: 10, loss: 0.20348145067691803
step: 20, loss: 0.55207359790802
step: 30, loss: 0.16168104112148285
step: 40, loss: 0.22812671959400177
step: 50, loss: 0.3374881148338318
step: 60, loss: 0.0942637100815773
step: 70, loss: 0.08285839110612869
step: 80, loss: 0.03840147331357002
step: 90, loss: 0.11000262945890427
step: 100, loss: 0.07948897033929825
step: 110, loss: 0.12755204737186432
step: 120, loss: 0.06675046682357788
step: 130, loss: 0.23816277086734772
step: 140, loss: 0.18314655125141144
step: 150, loss: 0.2747414708137512
step: 160, loss: 0.1691632866859436
step: 170, loss: 0.032424770295619965
step: 180, loss: 0.07063627988100052
step: 190, loss: 0.042619846761226654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7878787878787877, f1=0.8209366391184573,
best_f1=0.8209366391184573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14807549118995667
step: 10, loss: 0.1409541219472885
step: 20, loss: 0.20377618074417114
step: 30, loss: 0.040311846882104874
step: 40, loss: 0.07121063768863678
step: 50, loss: 0.08827845752239227
step: 60, loss: 0.12208797037601471
step: 70, loss: 0.238322913646698
step: 80, loss: 0.06904889643192291
step: 90, loss: 0.10343635082244873
step: 100, loss: 0.15540924668312073
step: 110, loss: 0.22846360504627228
step: 120, loss: 0.06633871793746948
step: 130, loss: 0.10484970360994339
step: 140, loss: 0.022373022511601448
step: 150, loss: 0.21142862737178802
step: 160, loss: 0.04656197130680084
step: 170, loss: 0.271311491727829
step: 180, loss: 0.13241751492023468
step: 190, loss: 0.0628742203116417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8324873096446701, f1=0.8250000000000001,
best_f1=0.8250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006606297101825476
step: 10, loss: 0.15072844922542572
step: 20, loss: 0.03158832713961601
step: 30, loss: 0.009727288037538528
step: 40, loss: 0.16497744619846344
step: 50, loss: 0.016582610085606575
step: 60, loss: 0.06049366295337677
step: 70, loss: 0.016124499961733818
step: 80, loss: 0.018519466742873192
step: 90, loss: 0.11472712457180023
step: 100, loss: 0.06862403452396393
step: 110, loss: 0.22485285997390747
step: 120, loss: 0.05353206396102905
step: 130, loss: 0.06866749376058578

```

```

step: 140, loss: 0.07322867214679718
step: 150, loss: 0.03405465930700302
step: 160, loss: 0.01311039924621582
step: 170, loss: 0.03193790093064308
step: 180, loss: 0.02840002439916134
step: 190, loss: 0.016257045790553093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8467532467532467, f1=0.8410256410256409,
best_f1=0.8410256410256409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1134011521935463
step: 10, loss: 0.10261248052120209
step: 20, loss: 0.02336614392697811
step: 30, loss: 0.0453641451895237
step: 40, loss: 0.0175747312605381
step: 50, loss: 0.02318798005580902
step: 60, loss: 0.003460885025560856
step: 70, loss: 0.14784500002861023
step: 80, loss: 0.03954392671585083
step: 90, loss: 0.051679421216249466
step: 100, loss: 0.17453442513942719
step: 110, loss: 0.15698589384555817
step: 120, loss: 0.04397638142108917
step: 130, loss: 0.19506262242794037
step: 140, loss: 0.01259447168558836
step: 150, loss: 0.04207554832100868
step: 160, loss: 0.005705162882804871
step: 170, loss: 0.0073339687660336494
step: 180, loss: 0.03575993701815605
step: 190, loss: 0.03459124639630318
/usr/local/lib/python3.7/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=0.8453608247422681, f1=0.8431876606683805,
best_f1=0.8410256410256409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23774829506874084
step: 10, loss: 0.03792508691549301

```

```

step: 20, loss: 0.005996422842144966
step: 30, loss: 0.10580416023731232
step: 40, loss: 0.013240307569503784
step: 50, loss: 0.09382642060518265
step: 60, loss: 0.030068958178162575
step: 70, loss: 0.0642947256565094
step: 80, loss: 0.010809276252985
step: 90, loss: 0.00499291718006134
step: 100, loss: 0.01587776467204094
step: 110, loss: 0.02007843554019928
step: 120, loss: 0.007984889671206474
step: 130, loss: 0.02817501127719879
step: 140, loss: 0.0029750331304967403
step: 150, loss: 0.012788339518010616
step: 160, loss: 0.031165970489382744
step: 170, loss: 0.2294861376285553
step: 180, loss: 0.08103501796722412
step: 190, loss: 0.05023960769176483
/usr/local/lib/python3.7/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=0.8402948402948403, f1=0.8175182481751825,
best_f1=0.8410256410256409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03292789310216904
step: 10, loss: 0.019235271960496902
step: 20, loss: 0.003452896373346448
step: 30, loss: 0.005420548375695944
step: 40, loss: 0.0028906858060508966
step: 50, loss: 0.005679419729858637
step: 60, loss: 0.08350932598114014
step: 70, loss: 0.0067189643159508705
step: 80, loss: 0.004637652076780796
step: 90, loss: 0.008725419640541077
step: 100, loss: 0.27689576148986816
step: 110, loss: 0.031365636736154556
step: 120, loss: 0.009490781463682652
step: 130, loss: 0.023414844647049904
step: 140, loss: 0.013612378388643265
step: 150, loss: 0.024483857676386833
step: 160, loss: 0.0639696791768074
step: 170, loss: 0.20950186252593994
step: 180, loss: 0.025579525157809258
step: 190, loss: 0.06490945816040039
/usr/local/lib/python3.7/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=0.8438356164383561, f1=0.8287292817679558,
best_f1=0.8410256410256409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16954822838306427
step: 10, loss: 0.013835547491908073
step: 20, loss: 0.004487285856157541
step: 30, loss: 0.03775588423013687
step: 40, loss: 0.10196208208799362
step: 50, loss: 0.011806693859398365
step: 60, loss: 0.02296675555408001
step: 70, loss: 0.0017406082479283214
step: 80, loss: 0.02375626750290394
step: 90, loss: 0.0033972454257309437
step: 100, loss: 0.0010595222702249885
step: 110, loss: 0.023205434903502464
step: 120, loss: 0.011872641742229462
step: 130, loss: 0.027486184611916542
step: 140, loss: 0.001083885319530964
step: 150, loss: 0.0076582166366279125
step: 160, loss: 0.062496040016412735
step: 170, loss: 0.001307222992181778
step: 180, loss: 0.02249598130583763
step: 190, loss: 0.003894971450790763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8658227848101266, f1=0.8346055979643766,
best_f1=0.8346055979643766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010850751772522926
step: 10, loss: 0.0008403781685046852
step: 20, loss: 0.0018455215031281114
step: 30, loss: 0.045730363577604294
step: 40, loss: 0.02717052958905697
step: 50, loss: 0.012268577702343464
step: 60, loss: 0.014000636525452137
step: 70, loss: 0.004759580362588167
step: 80, loss: 0.003416192252188921
step: 90, loss: 0.01524095144122839

```

```

step: 100, loss: 0.00790708139538765
step: 110, loss: 0.03486676514148712
step: 120, loss: 0.001260923920199275
step: 130, loss: 0.005214056000113487
step: 140, loss: 0.13518016040325165
step: 150, loss: 0.0006086950888857245
step: 160, loss: 0.010412508621811867
step: 170, loss: 0.06898289173841476
step: 180, loss: 0.013969079591333866
step: 190, loss: 0.000580532185267657
/usr/local/lib/python3.7/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=0.8615384615384616, f1=0.8512820512820511,
best_f1=0.8346055979643766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005136397667229176
step: 10, loss: 0.0013485521776601672
step: 20, loss: 0.12223430722951889
step: 30, loss: 0.009233292192220688
step: 40, loss: 0.005186343099921942
step: 50, loss: 0.012521202675998211
step: 60, loss: 0.005089366342872381
step: 70, loss: 0.0070794085040688515
step: 80, loss: 0.001315023866482079
step: 90, loss: 0.0007071010186336935
step: 100, loss: 0.009528622031211853
step: 110, loss: 0.0023147074971348047
step: 120, loss: 0.008009287528693676
step: 130, loss: 0.005539633799344301
step: 140, loss: 0.008261696435511112
step: 150, loss: 0.0009833653457462788
step: 160, loss: 0.0020093764178454876
step: 170, loss: 0.0029961413238197565
step: 180, loss: 0.0013798828003928065
step: 190, loss: 0.0005934092914685607
/usr/local/lib/python3.7/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=0.8481675392670156, f1=0.8346456692913385,
best_f1=0.8346055979643766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08907506614923477
step: 10, loss: 0.004938718397170305

```



```

step: 20, loss: 0.006268676836043596
step: 30, loss: 0.004341774620115757
step: 40, loss: 0.0005792583106085658
step: 50, loss: 0.000703522702679038
step: 60, loss: 0.002058574231341481
step: 70, loss: 0.002154655521735549
step: 80, loss: 0.0030241769272834063
step: 90, loss: 0.18099163472652435
step: 100, loss: 0.0029647306073457003
step: 110, loss: 0.21895797550678253
step: 120, loss: 0.0008439167286269367
step: 130, loss: 0.0017001023516058922
step: 140, loss: 0.001077570952475071
step: 150, loss: 0.005193511489778757
step: 160, loss: 0.006623271852731705
step: 170, loss: 0.0019566002301871777
step: 180, loss: 0.0023619006387889385
step: 190, loss: 0.004733049776405096
/usr/local/lib/python3.7/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=0.8617021276595744, f1=0.8571428571428571,
best_f1=0.8346055979643766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014004812575876713
step: 10, loss: 0.010324339382350445
step: 20, loss: 0.0024077913258224726
step: 30, loss: 0.0010425568325445056
step: 40, loss: 0.0008499196846969426
step: 50, loss: 0.006410757079720497
step: 60, loss: 0.0010836642468348145
step: 70, loss: 0.0010059112682938576
step: 80, loss: 0.011022687889635563
step: 90, loss: 0.11547860503196716
step: 100, loss: 0.010375989601016045
step: 110, loss: 0.004153123591095209
step: 120, loss: 0.004409075248986483
step: 130, loss: 0.004841488320380449
step: 140, loss: 0.0020520847756415606
step: 150, loss: 0.0019137706840410829
step: 160, loss: 0.0016875097062438726
step: 170, loss: 0.0016581827076151967
step: 180, loss: 0.026125077158212662
step: 190, loss: 0.0009832645300775766
/usr/local/lib/python3.7/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=0.8626373626373626, f1=0.8440860215053764,
best_f1=0.8346055979643766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038172451313585043
step: 10, loss: 0.0010701370192691684
step: 20, loss: 0.0005956644890829921
step: 30, loss: 0.019313644617795944
step: 40, loss: 0.003266284940764308
step: 50, loss: 0.002273698803037405
step: 60, loss: 0.007136298809200525
step: 70, loss: 0.023886071518063545
step: 80, loss: 0.0016984273679554462
step: 90, loss: 0.0044230311177670956
step: 100, loss: 0.0007146928692236543
step: 110, loss: 0.005427948664873838
step: 120, loss: 0.0012741668615490198
step: 130, loss: 0.032712243497371674
step: 140, loss: 0.00323260435834527
step: 150, loss: 0.032536040991544724
step: 160, loss: 0.02736111730337143
step: 170, loss: 0.0018476839177310467
step: 180, loss: 0.0012408512411639094
step: 190, loss: 0.1624165028333664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8702702702702703, f1=0.851063829787234,
best_f1=0.851063829787234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015207853866741061
step: 10, loss: 0.008800903335213661
step: 20, loss: 0.0034775277599692345
step: 30, loss: 0.0030900302808731794
step: 40, loss: 0.0018729633884504437
step: 50, loss: 0.001628075959160924
step: 60, loss: 0.0013895552838221192
step: 70, loss: 0.03311920911073685
step: 80, loss: 0.0009025527397170663
step: 90, loss: 0.0010175363859161735

```

```

step: 100, loss: 0.0008139042765833437
step: 110, loss: 0.000825723516754806
step: 120, loss: 0.02926507219672203
step: 130, loss: 0.0013115783222019672
step: 140, loss: 0.0024601027835160494
step: 150, loss: 0.007823208346962929
step: 160, loss: 0.002140308264642954
step: 170, loss: 0.10988499969244003
step: 180, loss: 0.0029505707789212465
step: 190, loss: 0.0014678919687867165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8716577540106952, f1=0.851063829787234,
best_f1=0.851063829787234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013507840922102332
step: 10, loss: 0.004930389579385519
step: 20, loss: 0.04248786345124245
step: 30, loss: 0.0009955300483852625
step: 40, loss: 0.0015274613397195935
step: 50, loss: 0.04496832191944122
step: 60, loss: 0.0012255239998921752
step: 70, loss: 0.0037725481670349836
step: 80, loss: 0.0010189312743023038
step: 90, loss: 0.003475236240774393
step: 100, loss: 0.0007268514018505812
step: 110, loss: 0.004252285696566105
step: 120, loss: 0.0008955759694799781
step: 130, loss: 0.01165989600121975
step: 140, loss: 0.0022333175875246525
step: 150, loss: 0.001213761861436069
step: 160, loss: 0.0017725727520883083
step: 170, loss: 0.11426667124032974
step: 180, loss: 0.005001382902264595
step: 190, loss: 0.0023441470693796873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8723404255319149, f1=0.8503937007874015,
best_f1=0.8503937007874015
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.")
2049it [00:12, 160.35it/s]
load_f1 = 0.828496042216359
real_f1 = 0.802030456852792
733it [00:00, 3414.37it/s]/usr/local/lib/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, 138.09it/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: 0.49946609139442444
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4460003077983856
step: 20, loss: 0.31552228331565857
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.4951237440109253
step: 40, loss: 0.5110642313957214
step: 50, loss: 0.34812670946121216
step: 60, loss: 0.593329668045044
step: 70, loss: 0.3443063497543335
step: 80, loss: 0.2230350226163864
step: 90, loss: 0.23304979503154755
step: 100, loss: 0.14551210403442383
step: 110, loss: 0.42427149415016174
step: 120, loss: 0.302602082490921
step: 130, loss: 0.3079533278942108
step: 140, loss: 0.35235944390296936
step: 150, loss: 0.3303760290145874
step: 160, loss: 0.375442236661911
step: 170, loss: 0.2939412593841553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.25602409638554213, f1=0.23493975903614459,
best_f1=0.23493975903614459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3060239255428314
step: 10, loss: 0.391052782535553
step: 20, loss: 0.28122732043266296
step: 30, loss: 0.2997487187385559
step: 40, loss: 0.05250668153166771
step: 50, loss: 0.43421727418899536
step: 60, loss: 0.17444103956222534
step: 70, loss: 0.48966190218925476
step: 80, loss: 0.2047383189201355
step: 90, loss: 0.22263537347316742
step: 100, loss: 0.4394131600856781
step: 110, loss: 0.2331237494945526
step: 120, loss: 0.10641408711671829
step: 130, loss: 0.46381786465644836
step: 140, loss: 0.25988340377807617
step: 150, loss: 0.1686955690383911
step: 160, loss: 0.24554666876792908
step: 170, loss: 0.07112332433462143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6538461538461539, f1=0.7113163972286374,
best_f1=0.7113163972286374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4057770371437073
step: 10, loss: 0.11942015588283539
step: 20, loss: 0.11302333325147629
step: 30, loss: 0.05367425084114075
step: 40, loss: 0.19977815449237823
step: 50, loss: 0.4453384578227997
step: 60, loss: 0.08775912970304489
step: 70, loss: 0.09211534261703491
step: 80, loss: 0.09601099789142609
step: 90, loss: 0.18741141259670258
```

```

step: 100, loss: 0.10264275223016739
step: 110, loss: 0.061018187552690506
step: 120, loss: 0.05623771995306015
step: 130, loss: 0.21531455218791962
step: 140, loss: 0.17141513526439667
step: 150, loss: 0.07382028549909592
step: 160, loss: 0.11787716299295425
step: 170, loss: 0.05996428057551384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8441247002398082, f1=0.8387096774193549,
best_f1=0.8387096774193549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18129587173461914
step: 10, loss: 0.14583858847618103
step: 20, loss: 0.020393304526805878
step: 30, loss: 0.21526215970516205
step: 40, loss: 0.08632971346378326
step: 50, loss: 0.06900180876255035
step: 60, loss: 0.0849747359752655
step: 70, loss: 0.006531845312565565
step: 80, loss: 0.2229403555393219
step: 90, loss: 0.24904373288154602
step: 100, loss: 0.1137167289853096
step: 110, loss: 0.22659435868263245
step: 120, loss: 0.251568466424942
step: 130, loss: 0.07534930109977722
step: 140, loss: 0.07050593197345734
step: 150, loss: 0.27209779620170593
step: 160, loss: 0.014714464545249939
step: 170, loss: 0.08692995458841324
/usr/local/lib/python3.7/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=0.8341968911917099, f1=0.8619854721549638,
best_f1=0.8387096774193549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033270254731178284
step: 10, loss: 0.05691690742969513

```



```

step: 20, loss: 0.07743775099515915
step: 30, loss: 0.1021154522895813
step: 40, loss: 0.12611830234527588
step: 50, loss: 0.03295990824699402
step: 60, loss: 0.028056038543581963
step: 70, loss: 0.14992372691631317
step: 80, loss: 0.005505230277776718
step: 90, loss: 0.06169498339295387
step: 100, loss: 0.015845518559217453
step: 110, loss: 0.0881456658244133
step: 120, loss: 0.0376942977309227
step: 130, loss: 0.003099533962085843
step: 140, loss: 0.03898169845342636
step: 150, loss: 0.16080862283706665
step: 160, loss: 0.06774799525737762
step: 170, loss: 0.006930406205356121
/usr/local/lib/python3.7/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=0.8382352941176471, f1=0.8665105386416863,
best_f1=0.8387096774193549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16569368541240692
step: 10, loss: 0.031049447134137154
step: 20, loss: 0.06394059956073761
step: 30, loss: 0.014181056059896946
step: 40, loss: 0.005738222040235996
step: 50, loss: 0.0042457603849470615
step: 60, loss: 0.17532722651958466
step: 70, loss: 0.011019226163625717
step: 80, loss: 0.051162730902433395
step: 90, loss: 0.09581347554922104
step: 100, loss: 0.009509225375950336
step: 110, loss: 0.09491778165102005
step: 120, loss: 0.031406763941049576
step: 130, loss: 0.018345270305871964
step: 140, loss: 0.025143539533019066
step: 150, loss: 0.016848640516400337
step: 160, loss: 0.040315937250852585
step: 170, loss: 0.004255135077983141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8613861386138614, f1=0.8857142857142858,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03706609085202217
step: 10, loss: 0.05117371678352356
step: 20, loss: 0.0036170873790979385
step: 30, loss: 0.0058785611763596535
step: 40, loss: 0.0016406916547566652
step: 50, loss: 0.003353378502652049
step: 60, loss: 0.12301677465438843
step: 70, loss: 0.022233212366700172
step: 80, loss: 0.10398542881011963
step: 90, loss: 0.11737051606178284
step: 100, loss: 0.0029439840000122786
step: 110, loss: 0.06905784457921982
step: 120, loss: 0.007355856709182262
step: 130, loss: 0.004817042965441942
step: 140, loss: 0.01427407469600439
step: 150, loss: 0.1136229932308197
step: 160, loss: 0.009456011466681957
step: 170, loss: 0.04433875158429146
/usr/local/lib/python3.7/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=0.8599508599508598, f1=0.8787185354691076,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021088361740112305
step: 10, loss: 0.0016475694719702005
step: 20, loss: 0.01897965371608734
step: 30, loss: 0.001816942123696208
step: 40, loss: 0.006017814390361309
step: 50, loss: 0.0023878472857177258
step: 60, loss: 0.007888310588896275
step: 70, loss: 0.01297303382307291
step: 80, loss: 0.001863026525825262
step: 90, loss: 0.013452376239001751
step: 100, loss: 0.004214192740619183
step: 110, loss: 0.05031215772032738
step: 120, loss: 0.04363098368048668
step: 130, loss: 0.005313974805176258
step: 140, loss: 0.02066221833229065
step: 150, loss: 0.007491250988095999

```

```

step: 160, loss: 0.007711922749876976
step: 170, loss: 0.15157634019851685
/usr/local/lib/python3.7/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=0.8585131894484411, f1=0.8853211009174312,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041529323905706406
step: 10, loss: 0.01210914645344019
step: 20, loss: 0.0020024869590997696
step: 30, loss: 0.021610809490084648
step: 40, loss: 0.010464792139828205
step: 50, loss: 0.006169936154037714
step: 60, loss: 0.055795542895793915
step: 70, loss: 0.005341699812561274
step: 80, loss: 0.0033103972673416138
step: 90, loss: 0.02923215925693512
step: 100, loss: 0.03891913965344429
step: 110, loss: 0.05626431107521057
step: 120, loss: 0.0030978734139353037
step: 130, loss: 0.004541064612567425
step: 140, loss: 0.010014357976615429
step: 150, loss: 0.16436731815338135
step: 160, loss: 0.005820299033075571
step: 170, loss: 0.010221111588180065
/usr/local/lib/python3.7/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=0.8536585365853658, f1=0.8904428904428905,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003948726691305637
step: 10, loss: 0.015476372092962265
step: 20, loss: 0.10263434797525406
step: 30, loss: 0.08737089484930038
step: 40, loss: 0.007709315977990627
step: 50, loss: 0.0112157566472888
step: 60, loss: 0.06925120949745178
step: 70, loss: 0.017042217776179314
step: 80, loss: 0.10497517138719559
step: 90, loss: 0.0009457790292799473
step: 100, loss: 0.009086810983717442
step: 110, loss: 0.019361091777682304

```

```

step: 120, loss: 0.003979798406362534
step: 130, loss: 0.006721248850226402
step: 140, loss: 0.008116515353322029
step: 150, loss: 0.0009776426013559103
step: 160, loss: 0.0007918372284621
step: 170, loss: 0.04363764077425003
/usr/local/lib/python3.7/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=0.8557213930348259, f1=0.8936170212765958,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10125502198934555
step: 10, loss: 0.04541637748479843
step: 20, loss: 0.0004347770009189844
step: 30, loss: 0.00214960053563118
step: 40, loss: 0.0005994869861751795
step: 50, loss: 0.0812992975115776
step: 60, loss: 0.019360387697815895
step: 70, loss: 0.0023275616113096476
step: 80, loss: 0.23317866027355194
step: 90, loss: 0.007725957781076431
step: 100, loss: 0.0008687604567967355
step: 110, loss: 0.0015267791459336877
step: 120, loss: 0.002492043189704418
step: 130, loss: 0.037670258432626724
step: 140, loss: 0.033796682953834534
step: 150, loss: 0.004091456066817045
step: 160, loss: 0.10319056361913681
step: 170, loss: 0.050654809921979904
/usr/local/lib/python3.7/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=0.855072463768116, f1=0.8796296296296297,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001149154151789844
step: 10, loss: 0.002230940153822303
step: 20, loss: 0.0008795582689344883
step: 30, loss: 0.10106529295444489
step: 40, loss: 0.004839261993765831
step: 50, loss: 0.0005329198902472854
step: 60, loss: 0.009531517513096333
step: 70, loss: 0.002116893883794546

```

```

step: 80, loss: 0.0003775588993448764
step: 90, loss: 0.05073815584182739
step: 100, loss: 0.0008317993488162756
step: 110, loss: 0.014575749635696411
step: 120, loss: 0.011007721535861492
step: 130, loss: 0.13818277418613434
step: 140, loss: 0.0005261553451418877
step: 150, loss: 0.00639225821942091
step: 160, loss: 0.004030193202197552
step: 170, loss: 0.071546271443367
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 12: dev_f1=0.8571428571428572, f1=0.8942307692307693,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009652639273554087
step: 10, loss: 0.0014192662201821804
step: 20, loss: 0.0006420629797503352
step: 30, loss: 0.02869126759469509
step: 40, loss: 0.0013549299910664558
step: 50, loss: 0.0585147887468338
step: 60, loss: 0.0016395867569372058
step: 70, loss: 0.1290174424648285
step: 80, loss: 0.0006612066063098609
step: 90, loss: 0.002383200451731682
step: 100, loss: 0.015455314889550209
step: 110, loss: 0.002011169446632266
step: 120, loss: 0.0023151554632931948
step: 130, loss: 0.005934881512075663
step: 140, loss: 0.04366222769021988
step: 150, loss: 0.0015595457516610622
step: 160, loss: 0.00032133751665242016
step: 170, loss: 0.0009133205167017877
/usr/local/lib/python3.7/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=0.8564356435643565, f1=0.8888888888888889,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005798729602247477
step: 10, loss: 0.021052179858088493
step: 20, loss: 0.1234443336725235

```

```

step: 30, loss: 0.023313721641898155
step: 40, loss: 0.0006051964010111988
step: 50, loss: 0.0008112230570986867
step: 60, loss: 0.09683764725923538
step: 70, loss: 0.007208350580185652
step: 80, loss: 0.008596554398536682
step: 90, loss: 0.09392668306827545
step: 100, loss: 0.00037516874726861715
step: 110, loss: 0.0325034074485302
step: 120, loss: 0.013412514701485634
step: 130, loss: 0.09725264459848404
step: 140, loss: 0.02289196103811264
step: 150, loss: 0.05001048371195793
step: 160, loss: 0.0018802284030243754
step: 170, loss: 0.027963129803538322
/usr/local/lib/python3.7/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=0.8578680203045685, f1=0.8992628992628993,
best_f1=0.8857142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024058423878159374
step: 10, loss: 0.008732417598366737
step: 20, loss: 0.001110076904296875
step: 30, loss: 0.001038194983266294
step: 40, loss: 0.0034080171026289463
step: 50, loss: 0.0029711551032960415
step: 60, loss: 0.003602443728595972
step: 70, loss: 0.0867844745516777
step: 80, loss: 0.000629759335424751
step: 90, loss: 0.002539178589358926
step: 100, loss: 0.07025711983442307
step: 110, loss: 0.0007462264620698988
step: 120, loss: 0.00021872638899367303
step: 130, loss: 0.000294489786028862
step: 140, loss: 0.010763326659798622
step: 150, loss: 0.0011824104003608227
step: 160, loss: 0.00032889406429603696
step: 170, loss: 0.0012654847232624888
/usr/local/lib/python3.7/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=0.8467153284671532, f1=0.8796296296296297,
best_f1=0.8857142857142858
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32

```

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

1916it [00:08, 227.83it/s]

load_f1 = 0.42815249266862176

real_f1 = 0.376068376068376

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, 137.96it/s]

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

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

```

--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm 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, 477kB/s]
 Downloading: 100% 899k/899k [00:00<00:00, 6.11MB/s]
 Downloading: 100% 456k/456k [00:00<00:00, 3.76MB/s]
 Downloading: 100% 501M/501M [00:09<00:00, 51.6MB/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: 0.5652621388435364
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.437966525554657
step: 20, loss: 0.4647291898727417
step: 30, loss: 0.32308825850486755
step: 40, loss: 0.34304824471473694
step: 50, loss: 0.5848415493965149
step: 60, loss: 0.5142657160758972
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.34145939350128174
step: 80, loss: 0.15235039591789246
step: 90, loss: 0.0939095988869667
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 100, loss: 0.14792372286319733
step: 110, loss: 0.26496824622154236
step: 120, loss: 0.06938249617815018
step: 130, loss: 0.12519918382167816
step: 140, loss: 0.12515726685523987
step: 150, loss: 0.4210265278816223
step: 160, loss: 0.016384519636631012
step: 170, loss: 0.11939571797847748
step: 180, loss: 0.13965871930122375
step: 190, loss: 0.07479539513587952
step: 200, loss: 0.010223706252872944
step: 210, loss: 0.024609621614217758
step: 220, loss: 0.03874139115214348
step: 230, loss: 0.04670604690909386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.957095709570957, f1=0.9634551495016611,
best_f1=0.9634551495016611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028104064986109734
step: 10, loss: 0.07670847326517105
step: 20, loss: 0.02519775554537773
step: 30, loss: 0.12720608711242676
step: 40, loss: 0.03841507434844971
step: 50, loss: 0.021349651739001274
step: 60, loss: 0.0044394005089998245
step: 70, loss: 0.011046959087252617
step: 80, loss: 0.0029299280140548944
step: 90, loss: 0.0028715587686747313
step: 100, loss: 0.0028231139294803143
step: 110, loss: 0.013890706934034824
step: 120, loss: 0.005515965633094311
step: 130, loss: 0.061183325946331024
step: 140, loss: 0.008165599778294563
step: 150, loss: 0.09773753583431244
step: 160, loss: 0.021577363833785057
step: 170, loss: 0.006097453646361828
step: 180, loss: 0.0052842628210783005
step: 190, loss: 0.10872990638017654
step: 200, loss: 0.050545237958431244
step: 210, loss: 0.01682932861149311
step: 220, loss: 0.018172139301896095
step: 230, loss: 0.00468862522393465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9732142857142857, 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: 0.0037213449832051992
step: 10, loss: 0.01643238216638565
step: 20, loss: 0.020275918766856194

```

```

step: 30, loss: 0.005058291833847761
step: 40, loss: 0.007009590044617653
step: 50, loss: 0.03204305097460747
step: 60, loss: 0.21657419204711914
step: 70, loss: 0.006410876754671335
step: 80, loss: 0.11907719075679779
step: 90, loss: 0.033914532512426376
step: 100, loss: 0.049521513283252716
step: 110, loss: 0.007394751999527216
step: 120, loss: 0.0026280004531145096
step: 130, loss: 0.004152992274612188
step: 140, loss: 0.01640208810567856
step: 150, loss: 0.01422099769115448
step: 160, loss: 0.004665189888328314
step: 170, loss: 0.0013701217249035835
step: 180, loss: 0.01737009920179844
step: 190, loss: 0.016750944778323174
step: 200, loss: 0.007805796805769205
step: 210, loss: 0.007886050269007683
step: 220, loss: 0.17007631063461304
step: 230, loss: 0.010290795005857944
/usr/local/lib/python3.7/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=0.9664429530201343, f1=0.967452300785634,
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: 0.06252933293581009
step: 10, loss: 0.0013580096419900656
step: 20, loss: 0.0026850756257772446
step: 30, loss: 0.005935355555266142
step: 40, loss: 0.1378570944070816
step: 50, loss: 0.07236552238464355
step: 60, loss: 0.01887674629688263
step: 70, loss: 0.007023169659078121
step: 80, loss: 0.10713966190814972
step: 90, loss: 0.03538786992430687
step: 100, loss: 0.01692374236881733
step: 110, loss: 0.006627555005252361
step: 120, loss: 0.015097656287252903
step: 130, loss: 0.003336122492328286
step: 140, loss: 0.0038073090836405754
step: 150, loss: 0.012348226271569729
step: 160, loss: 0.02228403277695179
step: 170, loss: 0.004859857726842165
step: 180, loss: 0.14669156074523926

```

```

step: 190, loss: 0.0028842128813266754
step: 200, loss: 0.11318991333246231
step: 210, loss: 0.0035623398143798113
step: 220, loss: 0.00898351613432169
step: 230, loss: 0.0036813104525208473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776785714285714, f1=0.9696969696969697,
best_f1=0.9696969696969697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021183995995670557
step: 10, loss: 0.004718332551419735
step: 20, loss: 0.037687163800001144
step: 30, loss: 0.0024462377186864614
step: 40, loss: 0.004680532496422529
step: 50, loss: 0.023427801206707954
step: 60, loss: 0.006646409630775452
step: 70, loss: 0.004840285051614046
step: 80, loss: 0.02574029564857483
step: 90, loss: 0.027309250086545944
step: 100, loss: 0.0006883603055030107
step: 110, loss: 0.0017600602004677057
step: 120, loss: 0.0007906926912255585
step: 130, loss: 0.0005725838127546012
step: 140, loss: 0.03250168263912201
step: 150, loss: 0.03138967230916023
step: 160, loss: 0.008205100893974304
step: 170, loss: 0.03752743825316429
step: 180, loss: 0.0036732531152665615
step: 190, loss: 0.04742709547281265
step: 200, loss: 0.19671232998371124
step: 210, loss: 0.005751245655119419
step: 220, loss: 0.004452534485608339
step: 230, loss: 0.19176989793777466
/usr/local/lib/python3.7/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=0.9690949227373068, f1=0.9668141592920354,
best_f1=0.9696969696969697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011972844367846847
step: 10, loss: 0.0037849145010113716
step: 20, loss: 0.003130035474896431
step: 30, loss: 0.004415353760123253
step: 40, loss: 0.00134156399872154
step: 50, loss: 0.0025620730593800545
step: 60, loss: 0.005944469477981329
step: 70, loss: 0.10756313800811768
step: 80, loss: 0.002213824773207307
step: 90, loss: 0.012605334632098675
step: 100, loss: 0.0011666283244267106
step: 110, loss: 0.007659317925572395
step: 120, loss: 0.003643251257017255
step: 130, loss: 0.002004149602726102
step: 140, loss: 0.0021704421378672123
step: 150, loss: 0.00020441258675418794
step: 160, loss: 0.0358358733355999
step: 170, loss: 0.005677691660821438
step: 180, loss: 0.00668680015951395
step: 190, loss: 0.0043038371950387955
step: 200, loss: 0.0024694546591490507
step: 210, loss: 0.0018109437078237534
step: 220, loss: 0.0013635840732604265
step: 230, loss: 0.0008981929277069867
/usr/local/lib/python3.7/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=0.975609756097561, f1=0.9720670391061451,
best_f1=0.9696969696969697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013784099137410522
step: 10, loss: 0.0016952133737504482
step: 20, loss: 0.004411640111356974
step: 30, loss: 0.0008729567634873092
step: 40, loss: 0.0018950594821944833
step: 50, loss: 0.0023804050870239735
step: 60, loss: 0.004994613584131002
step: 70, loss: 0.005699246656149626
step: 80, loss: 0.002198896836489439
step: 90, loss: 0.000504359370097518
step: 100, loss: 0.0007344610057771206
step: 110, loss: 0.0006180906784720719
step: 120, loss: 0.0077406009659171104
step: 130, loss: 0.0027146595530211926
step: 140, loss: 0.0002647603687364608

```

```

step: 150, loss: 0.14247047901153564
step: 160, loss: 0.001108209602534771
step: 170, loss: 0.001005613012239337
step: 180, loss: 0.001074346830137074
step: 190, loss: 0.07426200807094574
step: 200, loss: 0.004142874386161566
step: 210, loss: 0.011900652199983597
step: 220, loss: 0.0012069431832060218
step: 230, loss: 0.005216526798903942
/usr/local/lib/python3.7/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=0.9768467475192942, f1=0.9605263157894737,
best_f1=0.9696969696969697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00881531648337841
step: 10, loss: 0.0066000609658658504
step: 20, loss: 0.0025051524862647057
step: 30, loss: 0.01636127382516861
step: 40, loss: 0.002549991477280855
step: 50, loss: 0.008369065821170807
step: 60, loss: 0.0025840168818831444
step: 70, loss: 0.00025850083329714835
step: 80, loss: 0.0546397864818573
step: 90, loss: 0.0007056394242681563
step: 100, loss: 0.0005655837594531476
step: 110, loss: 0.005297339521348476
step: 120, loss: 0.06271930038928986
step: 130, loss: 0.011862711980938911
step: 140, loss: 0.0006465300102718174
step: 150, loss: 0.020151976495981216
step: 160, loss: 0.004895898513495922
step: 170, loss: 0.0042500803247094154
step: 180, loss: 0.0018723740940913558
step: 190, loss: 0.0008567770128138363
step: 200, loss: 0.010387251153588295
step: 210, loss: 0.002187473699450493
step: 220, loss: 0.004829758312553167
step: 230, loss: 0.0007040222990326583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832026875699889, f1=0.9765363128491621,
best_f1=0.9765363128491621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016903217183426023
step: 10, loss: 0.016309527680277824
step: 20, loss: 0.001126883435063064
step: 30, loss: 0.005805748514831066
step: 40, loss: 0.0067359330132603645
step: 50, loss: 0.00063103879801929
step: 60, loss: 0.0008943896973505616
step: 70, loss: 0.08125583827495575
step: 80, loss: 0.00028733417275361717
step: 90, loss: 0.022011183202266693
step: 100, loss: 0.0005165969487279654
step: 110, loss: 0.000371792382793501
step: 120, loss: 0.15173077583312988
step: 130, loss: 0.006730883382260799
step: 140, loss: 0.0018608418758958578
step: 150, loss: 0.0009333798661828041
step: 160, loss: 0.013481917791068554
step: 170, loss: 0.0004214798682369292
step: 180, loss: 0.004443150479346514
step: 190, loss: 0.0002627268258947879
step: 200, loss: 0.08182341605424881
step: 210, loss: 0.03667791187763214
step: 220, loss: 0.0006463530007749796
step: 230, loss: 0.00038561527617275715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854748603351955, 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: 0.0005555617390200496
step: 10, loss: 0.0008218658622354269
step: 20, loss: 0.0004271489742677659
step: 30, loss: 0.00045159977162256837
step: 40, loss: 0.0013506717514246702
step: 50, loss: 0.00043829932110384107
step: 60, loss: 0.0011897567892447114

```

```

step: 70, loss: 0.029550701379776
step: 80, loss: 0.0008167421910911798
step: 90, loss: 0.0004822936898563057
step: 100, loss: 0.0012761830585077405
step: 110, loss: 0.004044306930154562
step: 120, loss: 0.0010823324555531144
step: 130, loss: 0.0017816799227148294
step: 140, loss: 0.003097260370850563
step: 150, loss: 0.004905008710920811
step: 160, loss: 0.0001973260223167017
step: 170, loss: 0.00047960958909243345
step: 180, loss: 0.007817329838871956
step: 190, loss: 0.000574081321246922
step: 200, loss: 0.001163769280537963
step: 210, loss: 0.0008032607147470117
step: 220, loss: 0.001105845789425075
step: 230, loss: 0.00024152002879418433
/usr/local/lib/python3.7/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=0.9786276715410572, f1=0.963718820861678,
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: 0.0019839536398649216
step: 10, loss: 0.0005556613905355334
step: 20, loss: 0.000489825033582747
step: 30, loss: 0.00058103067567572
step: 40, loss: 9.885311737889424e-05
step: 50, loss: 0.00024018534168135375
step: 60, loss: 0.010355648584663868
step: 70, loss: 0.001270010368898511
step: 80, loss: 0.021632488816976547
step: 90, loss: 0.1937570869922638
step: 100, loss: 0.0006802348070777953
step: 110, loss: 0.0008542008581571281
step: 120, loss: 0.0006659352802671492
step: 130, loss: 0.0008935320074670017
step: 140, loss: 0.0011029712622985244
step: 150, loss: 0.0006827716715633869
step: 160, loss: 0.0008054712088778615
step: 170, loss: 0.0007156352512538433
step: 180, loss: 0.0010986004490405321
step: 190, loss: 0.0006144847138784826
step: 200, loss: 0.000908291491214186
step: 210, loss: 0.0009009927161969244
step: 220, loss: 0.0011780145578086376

```



```

step: 230, loss: 0.0015678582713007927
/usr/local/lib/python3.7/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=0.9809203142536477, 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: 0.0073250471614301205
step: 10, loss: 0.0005169233772903681
step: 20, loss: 0.010490653105080128
step: 30, loss: 0.03904583305120468
step: 40, loss: 0.0010067573748528957
step: 50, loss: 0.008408138528466225
step: 60, loss: 0.0009612705325707793
step: 70, loss: 0.0006873246748000383
step: 80, loss: 0.0005812391173094511
step: 90, loss: 0.009350378066301346
step: 100, loss: 0.00041373042040504515
step: 110, loss: 0.0003767125599551946
step: 120, loss: 0.009889715351164341
step: 130, loss: 0.0014976823003962636
step: 140, loss: 0.0006418749690055847
step: 150, loss: 0.0006809781771153212
step: 160, loss: 0.0010076271137222648
step: 170, loss: 0.0006346335867419839
step: 180, loss: 0.0005042356788180768
step: 190, loss: 0.0021599754691123962
step: 200, loss: 0.000448443868663162
step: 210, loss: 0.0010347474599257112
step: 220, loss: 0.0050199804827570915
step: 230, loss: 0.001111502991989255
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9708520179372198,
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: 0.002286649541929364
step: 10, loss: 0.0024889737833291292
step: 20, loss: 0.0010342577006667852
step: 30, loss: 0.0005139092681929469
step: 40, loss: 0.001440091640688479
step: 50, loss: 0.0016157529316842556
step: 60, loss: 0.000718731782399118

```

```

step: 70, loss: 0.0005857626674696803
step: 80, loss: 0.0004930120194330812
step: 90, loss: 0.0006379776750691235
step: 100, loss: 0.0013358454452827573
step: 110, loss: 0.003205278655514121
step: 120, loss: 0.0011002623941749334
step: 130, loss: 0.0011204509064555168
step: 140, loss: 0.0033004675060510635
step: 150, loss: 0.0003310564497951418
step: 160, loss: 0.004794502630829811
step: 170, loss: 0.0010742804734036326
step: 180, loss: 0.024443775415420532
step: 190, loss: 0.0009576123557053506
step: 200, loss: 0.0003143741050735116
step: 210, loss: 0.0008962743449956179
step: 220, loss: 0.0007908743573352695
step: 230, loss: 0.000873473531100899
/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9685393258426966,
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: 0.0005102249560877681
step: 10, loss: 0.00038538797525689006
step: 20, loss: 0.0034342824947088957
step: 30, loss: 0.000911604322027415
step: 40, loss: 0.0005879842210561037
step: 50, loss: 0.0005351973813958466
step: 60, loss: 0.0005888671730645001
step: 70, loss: 0.0004377181176096201
step: 80, loss: 0.0005992927472107112
step: 90, loss: 0.0010020730551332235
step: 100, loss: 0.0008027374860830605
step: 110, loss: 0.0012842464493587613
step: 120, loss: 0.00032246479531750083
step: 130, loss: 0.007095718290656805
step: 140, loss: 0.0007864629151299596
step: 150, loss: 0.00031226049759425223
step: 160, loss: 0.0006011967197991908
step: 170, loss: 0.0006950083188712597
step: 180, loss: 0.000794108142144978
step: 190, loss: 0.0004206410376355052
step: 200, loss: 0.0007040145574137568
step: 210, loss: 0.0003871329827234149
step: 220, loss: 0.0005013024783693254

```

```

step: 230, loss: 0.13593965768814087
/usr/local/lib/python3.7/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=0.983277591973244, f1=0.9731543624161074,
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: 0.016887249425053596
step: 10, loss: 0.0005719586042687297
step: 20, loss: 0.0005166164482943714
step: 30, loss: 0.000715899805072695
step: 40, loss: 0.00032602966530248523
step: 50, loss: 0.0005083020078018308
step: 60, loss: 0.0250502098351717
step: 70, loss: 0.00042486260645091534
step: 80, loss: 0.0005394461913965642
step: 90, loss: 0.00038245850009843707
step: 100, loss: 0.00033231571433134377
step: 110, loss: 0.0006106935907155275
step: 120, loss: 0.05126513913273811
step: 130, loss: 0.0003910586819984019
step: 140, loss: 0.0019581643864512444
step: 150, loss: 0.0007561034872196615
step: 160, loss: 0.002604454755783081
step: 170, loss: 0.00020587447215802968
step: 180, loss: 0.0005245626089163125
step: 190, loss: 0.0015242320951074362
step: 200, loss: 0.03186763450503349
step: 210, loss: 0.0318855419754982
step: 220, loss: 0.0005050910986028612
step: 230, loss: 0.0007830127724446356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865771812080537, f1=0.9730337078651685,
best_f1=0.9730337078651685
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled          : True

```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 217.31it/s]
load_f1 = 0.9854096520763187
real_f1 = 0.9832402234636871
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.55it/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/

```

```

Downloading: 100% 481/481 [00:00<00:00, 430kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 36.5MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 29.6MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 66.4MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6192041635513306
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: 0.42061594128608704
step: 20, loss: 0.22992293536663055
step: 30, loss: 0.24529311060905457
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.2390555739402771
step: 50, loss: 0.20267736911773682
step: 60, loss: 0.1320333033800125
step: 70, loss: 0.23522938787937164
step: 80, loss: 0.1436021476984024
step: 90, loss: 0.11438119411468506
step: 100, loss: 0.14141888916492462
step: 110, loss: 0.1074221134185791
step: 120, loss: 0.07349252700805664
step: 130, loss: 0.12761618196964264
step: 140, loss: 0.25035884976387024
step: 150, loss: 0.09899695217609406
step: 160, loss: 0.13672205805778503
step: 170, loss: 0.05136435106396675
step: 180, loss: 0.06618360430002213
step: 190, loss: 0.03888415917754173
step: 200, loss: 0.029499946162104607
step: 210, loss: 0.056892912834882736
step: 220, loss: 0.04483944550156593
step: 230, loss: 0.34760963916778564
step: 240, loss: 0.05747292563319206
step: 250, loss: 0.03763753920793533
step: 260, loss: 0.06153273209929466
step: 270, loss: 0.37876245379447937
step: 280, loss: 0.058497726917266846
step: 290, loss: 0.0781528428196907
step: 300, loss: 0.23473413288593292
step: 310, loss: 0.08740942180156708
step: 320, loss: 0.12462839484214783
step: 330, loss: 0.056550413370132446
step: 340, loss: 0.49884986877441406
step: 350, loss: 0.07575032114982605
step: 360, loss: 0.10805746167898178
step: 370, loss: 0.14091849327087402
step: 380, loss: 0.19941478967666626
step: 390, loss: 0.04772931709885597
step: 400, loss: 0.07421204447746277
step: 410, loss: 0.29901790618896484
step: 420, loss: 0.024589620530605316
```

```

step: 430, loss: 0.033210638910532
step: 440, loss: 0.1210150346159935
step: 450, loss: 0.028244933113455772
step: 460, loss: 0.026181660592556
step: 470, loss: 0.025911778211593628
step: 480, loss: 0.22207759320735931
step: 490, loss: 0.16373570263385773
step: 500, loss: 0.05508720874786377
step: 510, loss: 0.0420939140021801
step: 520, loss: 0.11098875105381012
step: 530, loss: 0.009263833984732628
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9471750114836931, f1=0.9445208619899129,
best_f1=0.9445208619899129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05400262773036957
step: 10, loss: 0.04620395228266716
step: 20, loss: 0.04672104865312576
step: 30, loss: 0.02236497588455677
step: 40, loss: 0.035796817392110825
step: 50, loss: 0.06443820148706436
step: 60, loss: 0.01455804705619812
step: 70, loss: 0.009738834574818611
step: 80, loss: 0.02098110131919384
step: 90, loss: 0.009652435779571533
step: 100, loss: 0.07843751460313797
step: 110, loss: 0.00480164960026741
step: 120, loss: 0.10302302241325378
step: 130, loss: 0.017770541831851006
step: 140, loss: 0.03592377528548241
step: 150, loss: 0.0274515338242054
step: 160, loss: 0.03685346990823746
step: 170, loss: 0.016748545691370964
step: 180, loss: 0.025723261758685112
step: 190, loss: 0.016762495040893555
step: 200, loss: 0.25227949023246765
step: 210, loss: 0.01544070802628994
step: 220, loss: 0.004844301380217075
step: 230, loss: 0.10932447016239166
step: 240, loss: 0.06876280903816223

```

```

step: 250, loss: 0.011206898838281631
step: 260, loss: 0.09147952497005463
step: 270, loss: 0.036122336983680725
step: 280, loss: 0.09949609637260437
step: 290, loss: 0.044989679008722305
step: 300, loss: 0.03455604612827301
step: 310, loss: 0.15701629221439362
step: 320, loss: 0.0501287467777729
step: 330, loss: 0.02423129603266716
step: 340, loss: 0.1280559003353119
step: 350, loss: 0.007638570386916399
step: 360, loss: 0.03280983865261078
step: 370, loss: 0.0039217970333993435
step: 380, loss: 0.10768625885248184
step: 390, loss: 0.014038597233593464
step: 400, loss: 0.023979108780622482
step: 410, loss: 0.01975334621965885
step: 420, loss: 0.06011027842760086
step: 430, loss: 0.18498504161834717
step: 440, loss: 0.004493745509535074
step: 450, loss: 0.037325188517570496
step: 460, loss: 0.04866146296262741
step: 470, loss: 0.028623508289456367
step: 480, loss: 0.027199480682611465
step: 490, loss: 0.015247704461216927
step: 500, loss: 0.0030074219685047865
step: 510, loss: 0.022423017770051956
step: 520, loss: 0.4090724587440491
step: 530, loss: 0.09906478971242905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9503868912152936, f1=0.9448244414044686,
best_f1=0.9448244414044686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23229709267616272
step: 10, loss: 0.2527567148208618
step: 20, loss: 0.05231362208724022
step: 30, loss: 0.06283887475728989
step: 40, loss: 0.08083005249500275
step: 50, loss: 0.014586787670850754
step: 60, loss: 0.02185680717229843

```


step: 70, loss: 0.00481470488011837
step: 80, loss: 0.11391264945268631
step: 90, loss: 0.02093466930091381
step: 100, loss: 0.05865269526839256
step: 110, loss: 0.07892369478940964
step: 120, loss: 0.20023463666439056
step: 130, loss: 0.05030664801597595
step: 140, loss: 0.004789766389876604
step: 150, loss: 0.034751761704683304
step: 160, loss: 0.01293387170881033
step: 170, loss: 0.007799673359841108
step: 180, loss: 0.03262082859873772
step: 190, loss: 0.009870133362710476
step: 200, loss: 0.021404560655355453
step: 210, loss: 0.06789317727088928
step: 220, loss: 0.052752964198589325
step: 230, loss: 0.07524017244577408
step: 240, loss: 0.06253323704004288
step: 250, loss: 0.028652897104620934
step: 260, loss: 0.15468813478946686
step: 270, loss: 0.001487922971136868
step: 280, loss: 0.025031432509422302
step: 290, loss: 0.008597292937338352
step: 300, loss: 0.05107836052775383
step: 310, loss: 0.07047045975923538
step: 320, loss: 0.041727934032678604
step: 330, loss: 0.019360654056072235
step: 340, loss: 0.018819065764546394
step: 350, loss: 0.0917888954281807
step: 360, loss: 0.021425144746899605
step: 370, loss: 0.030032401904463768
step: 380, loss: 0.00969228707253933
step: 390, loss: 0.0029355615843087435
step: 400, loss: 0.10447905212640762
step: 410, loss: 0.014476859010756016
step: 420, loss: 0.019427213817834854
step: 430, loss: 0.028094664216041565
step: 440, loss: 0.14412759244441986
step: 450, loss: 0.04410263150930405
step: 460, loss: 0.012332110665738583
step: 470, loss: 0.08696486800909042
step: 480, loss: 0.32713788747787476
step: 490, loss: 0.04911962151527405
step: 500, loss: 0.005360263399779797
step: 510, loss: 0.030561288818717003
step: 520, loss: 0.009840602055191994
step: 530, loss: 0.0030387816950678825

/usr/local/lib/python3.7/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=0.9484346224677717, f1=0.9444699403396054,
best_f1=0.9448244414044686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009566509164869785
step: 10, loss: 0.0036902374122291803
step: 20, loss: 0.012039478868246078
step: 30, loss: 0.14297173917293549
step: 40, loss: 0.008099122904241085
step: 50, loss: 0.03868883103132248
step: 60, loss: 0.0012309482553973794
step: 70, loss: 0.042794980108737946
step: 80, loss: 0.0326785184442997
step: 90, loss: 0.21461625397205353
step: 100, loss: 0.00184333766810596
step: 110, loss: 0.104461170732975
step: 120, loss: 0.0045614587143063545
step: 130, loss: 0.11068753153085709
step: 140, loss: 0.04170847684144974
step: 150, loss: 0.013119440525770187
step: 160, loss: 0.022190580144524574
step: 170, loss: 0.007574389688670635
step: 180, loss: 0.07698559015989304
step: 190, loss: 0.0541035458445549
step: 200, loss: 0.012642348185181618
step: 210, loss: 0.0010660601546987891
step: 220, loss: 0.021599633619189262
step: 230, loss: 0.030026530846953392
step: 240, loss: 0.022111965343356133
step: 250, loss: 0.2455320507287979
step: 260, loss: 0.015885984525084496
step: 270, loss: 0.041009943932294846
step: 280, loss: 0.008871800266206264
step: 290, loss: 0.030257871374487877
step: 300, loss: 0.0016748631605878472
step: 310, loss: 0.01666482351720333
step: 320, loss: 0.05481318384408951
step: 330, loss: 0.028538601472973824
step: 340, loss: 0.011624799109995365
step: 350, loss: 0.19160041213035583
step: 360, loss: 0.09863916039466858
step: 370, loss: 0.01159665547311306
step: 380, loss: 0.026970939710736275
step: 390, loss: 0.0005016874056309462
step: 400, loss: 0.004290570504963398
```

```

step: 410, loss: 0.0037541647907346487
step: 420, loss: 0.015681428834795952
step: 430, loss: 0.008164887316524982
step: 440, loss: 0.02346203289926052
step: 450, loss: 0.001830823253840208
step: 460, loss: 0.04682070016860962
step: 470, loss: 0.008547765202820301
step: 480, loss: 0.23965132236480713
step: 490, loss: 0.020514996722340584
step: 500, loss: 0.04626990109682083
step: 510, loss: 0.050874143838882446
step: 520, loss: 0.015213815495371819
step: 530, loss: 0.12731730937957764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9531757070004635, f1=0.9478060046189377,
best_f1=0.9478060046189377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005925996811129153
step: 10, loss: 0.02620040811598301
step: 20, loss: 0.004097991622984409
step: 30, loss: 0.03542898967862129
step: 40, loss: 0.024603772908449173
step: 50, loss: 0.04452310502529144
step: 60, loss: 0.01910848543047905
step: 70, loss: 0.002129028784111142
step: 80, loss: 0.015407504513859749
step: 90, loss: 0.11283039301633835
step: 100, loss: 0.04663253203034401
step: 110, loss: 0.004996664822101593
step: 120, loss: 0.026900358498096466
step: 130, loss: 0.006130315829068422
step: 140, loss: 0.01282496191561222
step: 150, loss: 0.024426614865660667
step: 160, loss: 0.021440081298351288
step: 170, loss: 0.1270993947982788
step: 180, loss: 0.045855287462472916
step: 190, loss: 0.013482077047228813
step: 200, loss: 0.009325066581368446
step: 210, loss: 0.0024787727743387222
step: 220, loss: 0.0024505469482392073

```

```

step: 230, loss: 0.006341109983623028
step: 240, loss: 0.0031161399092525244
step: 250, loss: 0.08268766850233078
step: 260, loss: 0.0037703353445976973
step: 270, loss: 0.007707326672971249
step: 280, loss: 0.006842868868261576
step: 290, loss: 0.08098834753036499
step: 300, loss: 0.059420064091682434
step: 310, loss: 0.05001886188983917
step: 320, loss: 0.020871736109256744
step: 330, loss: 0.006304222624748945
step: 340, loss: 0.08565570414066315
step: 350, loss: 0.018590349704027176
step: 360, loss: 0.0017869811272248626
step: 370, loss: 0.0019291468197479844
step: 380, loss: 0.002299762796610594
step: 390, loss: 0.03132323920726776
step: 400, loss: 0.018258405849337578
step: 410, loss: 0.07204937934875488
step: 420, loss: 0.19684596359729767
step: 430, loss: 0.12732918560504913
step: 440, loss: 0.03629252687096596
step: 450, loss: 0.0011771711288020015
step: 460, loss: 0.0013976856134831905
step: 470, loss: 0.07716839015483856
step: 480, loss: 0.0021302548702806234
step: 490, loss: 0.015243025496602058
step: 500, loss: 0.004808025900274515
step: 510, loss: 0.011262129992246628
step: 520, loss: 0.23414729535579681
step: 530, loss: 0.03904463350772858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.953305594082293, f1=0.9448307834955958,
best_f1=0.9448307834955958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0672791451215744
step: 10, loss: 0.0058652665466070175
step: 20, loss: 0.038826629519462585
step: 30, loss: 0.008314820937812328
step: 40, loss: 0.03277435898780823

```

step: 50, loss: 0.037244927138090134
step: 60, loss: 0.03030168078839779
step: 70, loss: 0.00518877524882555
step: 80, loss: 0.0008582556620240211
step: 90, loss: 0.02023288980126381
step: 100, loss: 0.02172294817864895
step: 110, loss: 0.0037611224688589573
step: 120, loss: 0.023598941043019295
step: 130, loss: 0.0234824251383543
step: 140, loss: 0.0011433473555371165
step: 150, loss: 0.0030255145393311977
step: 160, loss: 0.057476483285427094
step: 170, loss: 0.0023591876961290836
step: 180, loss: 0.009619728662073612
step: 190, loss: 0.09210166335105896
step: 200, loss: 0.01168772391974926
step: 210, loss: 0.006843066308647394
step: 220, loss: 0.004800246097147465
step: 230, loss: 0.022611791267991066
step: 240, loss: 0.021872520446777344
step: 250, loss: 0.02458880841732025
step: 260, loss: 0.001602842123247683
step: 270, loss: 0.024826306849718094
step: 280, loss: 0.005881180055439472
step: 290, loss: 0.0025138144847005606
step: 300, loss: 0.0009352273191325366
step: 310, loss: 0.03289181366562843
step: 320, loss: 0.00012461523874662817
step: 330, loss: 0.00013479347398970276
step: 340, loss: 0.004647711757570505
step: 350, loss: 0.04588063806295395
step: 360, loss: 0.07725217938423157
step: 370, loss: 0.008197609335184097
step: 380, loss: 0.0002821333473548293
step: 390, loss: 0.001494846772402525
step: 400, loss: 0.002077801153063774
step: 410, loss: 0.0007359233568422496
step: 420, loss: 0.015101340599358082
step: 430, loss: 0.0012571396073326468
step: 440, loss: 0.00035111611941829324
step: 450, loss: 0.20145365595817566
step: 460, loss: 0.002279088832437992
step: 470, loss: 0.006308440119028091
step: 480, loss: 0.002802028553560376
step: 490, loss: 0.020890668034553528
step: 500, loss: 0.013188427314162254
step: 510, loss: 0.044258732348680496
step: 520, loss: 0.010350536555051804

```

step: 530, loss: 0.07635249942541122
/usr/local/lib/python3.7/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=0.9519450800915331, f1=0.9503219871205152,
best_f1=0.9448307834955958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0061426968313753605
step: 10, loss: 0.004266439937055111
step: 20, loss: 0.003759767860174179
step: 30, loss: 0.027372779324650764
step: 40, loss: 0.002611649688333273
step: 50, loss: 0.028255771845579147
step: 60, loss: 0.04215776547789574
step: 70, loss: 0.007419689558446407
step: 80, loss: 0.0017120620468631387
step: 90, loss: 4.3487074435688555e-05
step: 100, loss: 0.0012128258822485805
step: 110, loss: 0.0006109913229010999
step: 120, loss: 0.00038248844793997705
step: 130, loss: 0.00010153753100894392
step: 140, loss: 0.023281507194042206
step: 150, loss: 0.0015475377440452576
step: 160, loss: 0.0007058953051455319
step: 170, loss: 0.006659770384430885
step: 180, loss: 0.04000748321413994
step: 190, loss: 0.009274981915950775
step: 200, loss: 0.00029133621137589216
step: 210, loss: 0.028660379350185394
step: 220, loss: 0.0004445748927537352
step: 230, loss: 0.00023769262770656496
step: 240, loss: 0.012738481163978577
step: 250, loss: 0.006750416476279497
step: 260, loss: 0.008350305259227753
step: 270, loss: 0.009884574450552464
step: 280, loss: 0.019650902599096298
step: 290, loss: 0.027132641524076462
step: 300, loss: 0.0002610248629935086
step: 310, loss: 0.0018680975772440434
step: 320, loss: 0.012590198777616024
step: 330, loss: 0.05034756287932396
step: 340, loss: 0.00022569172142539173
step: 350, loss: 0.0008931146585382521
step: 360, loss: 0.05145176500082016
step: 370, loss: 0.061623990535736084
step: 380, loss: 0.014952531084418297

```

```

step: 390, loss: 0.028652707114815712
step: 400, loss: 0.009583046659827232
step: 410, loss: 0.07676564902067184
step: 420, loss: 0.0057695647701621056
step: 430, loss: 0.0016316954279318452
step: 440, loss: 0.0013343481114134192
step: 450, loss: 0.0021077743731439114
step: 460, loss: 0.0022332367952913046
step: 470, loss: 0.13045573234558105
step: 480, loss: 0.0069595081731677055
step: 490, loss: 0.05228304862976074
step: 500, loss: 0.003831644309684634
step: 510, loss: 0.0006332880584523082
step: 520, loss: 0.003225037595257163
step: 530, loss: 0.0014097014209255576
/usr/local/lib/python3.7/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=0.9508045977011494, f1=0.9487179487179488,
best_f1=0.9448307834955958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018294691108167171
step: 10, loss: 0.002543768147006631
step: 20, loss: 0.004808119032531977
step: 30, loss: 0.039153896272182465
step: 40, loss: 0.00292814034037292
step: 50, loss: 0.0012991157127544284
step: 60, loss: 0.0009377097594551742
step: 70, loss: 0.0016361345769837499
step: 80, loss: 0.0016114688478410244
step: 90, loss: 0.002230269368737936
step: 100, loss: 0.001613069325685501
step: 110, loss: 0.00024114648113027215
step: 120, loss: 0.00910047348588705
step: 130, loss: 0.0005741696804761887
step: 140, loss: 0.08456841856241226
step: 150, loss: 0.00015471488586626947
step: 160, loss: 0.00016384111950173974
step: 170, loss: 0.251120388507843
step: 180, loss: 0.05389935150742531
step: 190, loss: 0.13015352189540863
step: 200, loss: 0.002529274672269821
step: 210, loss: 0.11696470528841019
step: 220, loss: 0.0006167491083033383
step: 230, loss: 0.04729396104812622
step: 240, loss: 0.049099959433078766

```

```

step: 250, loss: 0.0006912434473633766
step: 260, loss: 0.09907716512680054
step: 270, loss: 0.003923804499208927
step: 280, loss: 0.0018727510469034314
step: 290, loss: 0.0005567530170083046
step: 300, loss: 6.7207969550509e-05
step: 310, loss: 0.0001953638275153935
step: 320, loss: 0.0024631766136735678
step: 330, loss: 8.598537533544004e-05
step: 340, loss: 0.014620751142501831
step: 350, loss: 0.0001776507415343076
step: 360, loss: 0.04283237084746361
step: 370, loss: 0.0920926183462143
step: 380, loss: 0.0003231749578844756
step: 390, loss: 0.033298309892416
step: 400, loss: 0.00024272105656564236
step: 410, loss: 0.0026894297916442156
step: 420, loss: 0.0010923744412139058
step: 430, loss: 0.03673752397298813
step: 440, loss: 0.01686405949294567
step: 450, loss: 0.0033804767299443483
step: 460, loss: 0.0008945664740167558
step: 470, loss: 0.05585227161645889
step: 480, loss: 0.0008553607040084898
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 490, loss: 0.004296059720218182
step: 500, loss: 0.00721860071644187
step: 510, loss: 0.017254862934350967
step: 520, loss: 0.0019321631407365203
step: 530, loss: 0.03312857821583748
/usr/local/lib/python3.7/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=0.9496738117427772, f1=0.9515828677839852,
best_f1=0.9448307834955958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014099841937422752
step: 10, loss: 0.010831667110323906
step: 20, loss: 0.0006582133355550468
step: 30, loss: 0.11252788454294205
step: 40, loss: 0.00032443689997307956
step: 50, loss: 0.0008417726494371891
step: 60, loss: 0.004725680686533451
step: 70, loss: 0.0018241427605971694
step: 80, loss: 0.005728122778236866
step: 90, loss: 0.086188405752182

```



```
step: 100, loss: 0.001690173172391951
step: 110, loss: 0.003993288613855839
step: 120, loss: 0.011425250209867954
step: 130, loss: 0.009232142008841038
step: 140, loss: 0.0012152419658377767
step: 150, loss: 0.0003231454757042229
step: 160, loss: 0.0006464723846875131
step: 170, loss: 0.0026145894080400467
step: 180, loss: 0.002084990730509162
step: 190, loss: 0.00013989255239721388
step: 200, loss: 0.0004528080753516406
step: 210, loss: 0.004132693167775869
step: 220, loss: 0.0010283185401931405
step: 230, loss: 0.00028794354875572026
step: 240, loss: 0.004633013159036636
step: 250, loss: 0.0010660062544047832
step: 260, loss: 0.0031508184038102627
step: 270, loss: 0.003624175675213337
step: 280, loss: 0.02726120315492153
step: 290, loss: 0.0013759013963863254
step: 300, loss: 0.0003974648134317249
step: 310, loss: 0.07775617390871048
step: 320, loss: 0.0011956709204241633
step: 330, loss: 0.0019680673722177744
step: 340, loss: 0.16725493967533112
step: 350, loss: 0.04481377825140953
step: 360, loss: 0.004569814540445805
step: 370, loss: 0.00028990284772589803
step: 380, loss: 0.0034243096597492695
step: 390, loss: 0.012470882385969162
step: 400, loss: 0.038779109716415405
step: 410, loss: 0.005488621536642313
step: 420, loss: 0.0004923614906147122
step: 430, loss: 0.009924665093421936
step: 440, loss: 0.004795448388904333
step: 450, loss: 0.0047142356634140015
step: 460, loss: 6.716937059536576e-05
step: 470, loss: 0.0008030221797525883
step: 480, loss: 0.00035194845986552536
step: 490, loss: 0.01584521494805813
step: 500, loss: 0.0008993318188004196
step: 510, loss: 0.0006594933802261949
step: 520, loss: 0.001192244584672153
step: 530, loss: 0.13572010397911072
/usr/local/lib/python3.7/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=0.9524253731343284, f1=0.9535315985130112,
```

```
best_f1=0.9448307834955958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015313275798689574
step: 10, loss: 0.001845119521021843
step: 20, loss: 0.00999737810343504
step: 30, loss: 0.014925702475011349
step: 40, loss: 0.00031215266790241003
step: 50, loss: 0.00027710371068678796
step: 60, loss: 0.0006178554031066597
step: 70, loss: 0.0006153436261229217
step: 80, loss: 0.0002412713220110163
step: 90, loss: 0.000501478963997215
step: 100, loss: 0.039647430181503296
step: 110, loss: 0.0032048108987510204
step: 120, loss: 0.09322979301214218
step: 130, loss: 0.005668903235346079
step: 140, loss: 0.005160126835107803
step: 150, loss: 0.0039035885129123926
step: 160, loss: 0.025228898972272873
step: 170, loss: 0.0003028438368346542
step: 180, loss: 0.007484239060431719
step: 190, loss: 0.0017552792560309172
step: 200, loss: 0.021183663979172707
step: 210, loss: 0.05372775346040726
step: 220, loss: 0.0030564030166715384
step: 230, loss: 0.0009476375998929143
step: 240, loss: 0.0012490934459492564
step: 250, loss: 0.001401794026605785
step: 260, loss: 0.0009893032256513834
step: 270, loss: 0.0007879786426201463
step: 280, loss: 0.02583528682589531
step: 290, loss: 0.001263403450138867
step: 300, loss: 0.005519193597137928
step: 310, loss: 0.03271261602640152
step: 320, loss: 0.0037468557711690664
step: 330, loss: 0.000868115748744458
step: 340, loss: 0.0003830556233879179
step: 350, loss: 0.22826290130615234
step: 360, loss: 0.002861594781279564
step: 370, loss: 0.0049677398055791855
step: 380, loss: 0.034075770527124405
step: 390, loss: 0.0012609590776264668
step: 400, loss: 0.0016882114578038454
step: 410, loss: 0.0014923728303983808
step: 420, loss: 0.0003671094309538603
step: 430, loss: 8.496589725837111e-05
```

```

step: 440, loss: 4.3055002606706694e-05
step: 450, loss: 0.0006635967292822897
step: 460, loss: 0.0023754097055643797
step: 470, loss: 0.0016129652503877878
step: 480, loss: 0.00011634452675934881
step: 490, loss: 0.0154905766248703
step: 500, loss: 0.002456237794831395
step: 510, loss: 0.0007272448856383562
step: 520, loss: 0.22871243953704834
step: 530, loss: 0.002597854472696781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9564007421150279, f1=0.9464867380176826,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006358437240123749
step: 10, loss: 0.004457168281078339
step: 20, loss: 0.0026170574128627777
step: 30, loss: 0.0003332838823553175
step: 40, loss: 8.814553439151496e-05
step: 50, loss: 0.0003214295720681548
step: 60, loss: 0.007030934561043978
step: 70, loss: 0.004005130846053362
step: 80, loss: 0.0007042540237307549
step: 90, loss: 0.0018199406331405044
step: 100, loss: 0.0016617703950032592
step: 110, loss: 0.00011222265311516821
step: 120, loss: 0.0008198143914341927
step: 130, loss: 0.004536938853561878
step: 140, loss: 0.00831600185483694
step: 150, loss: 0.00034760829294100404
step: 160, loss: 0.004104966763406992
step: 170, loss: 0.046015415340662
step: 180, loss: 0.0019930738490074873
step: 190, loss: 0.0015194466104730964
step: 200, loss: 0.004014225676655769
step: 210, loss: 0.002571044722571969
step: 220, loss: 0.008398675359785557
step: 230, loss: 0.0006425027386285365
step: 240, loss: 0.0023837294429540634
step: 250, loss: 0.0005756234750151634

```

```

step: 260, loss: 0.004655550699681044
step: 270, loss: 0.0028424605261534452
step: 280, loss: 0.0007033019210211933
step: 290, loss: 0.0015165593940764666
step: 300, loss: 0.0009445803589187562
step: 310, loss: 0.049925416707992554
step: 320, loss: 0.0014137611724436283
step: 330, loss: 7.495228055631742e-05
step: 340, loss: 0.009899014607071877
step: 350, loss: 0.021996309980750084
step: 360, loss: 0.0016310613136738539
step: 370, loss: 0.0015115069691091776
step: 380, loss: 0.00042125454638153315
step: 390, loss: 0.0010317808482795954
step: 400, loss: 0.000463821372250095
step: 410, loss: 0.001209342386573553
step: 420, loss: 0.0003063537005800754
step: 430, loss: 0.005706354044377804
step: 440, loss: 0.0004493450396694243
step: 450, loss: 0.010059588588774204
step: 460, loss: 0.0007650057668797672
step: 470, loss: 0.0002685084182303399
step: 480, loss: 0.0011272137053310871
step: 490, loss: 0.00010904627561103553
step: 500, loss: 0.002057305769994855
step: 510, loss: 6.283148832153529e-05
step: 520, loss: 0.0005765539244748652
step: 530, loss: 0.0007146972347982228
/usr/local/lib/python3.7/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=0.9513002364066194, f1=0.9469194312796209,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001770609524101019
step: 10, loss: 0.00035094801569357514
step: 20, loss: 0.008287382312119007
step: 30, loss: 1.687161966401618e-05
step: 40, loss: 0.0012874228414148092
step: 50, loss: 9.486715862294659e-05
step: 60, loss: 3.0367498766281642e-05
step: 70, loss: 0.0003976384177803993
step: 80, loss: 0.022406116127967834
step: 90, loss: 0.02914804220199585
step: 100, loss: 0.34992265701293945
step: 110, loss: 0.04714445024728775

```

```
step: 120, loss: 0.0008701742626726627
step: 130, loss: 0.0020166237372905016
step: 140, loss: 0.05582502856850624
step: 150, loss: 0.0030099956784397364
step: 160, loss: 0.06060980260372162
step: 170, loss: 0.0012284207623451948
step: 180, loss: 7.944636308820918e-05
step: 190, loss: 0.001056738430634141
step: 200, loss: 0.001966482726857066
step: 210, loss: 0.00020970906189177185
step: 220, loss: 6.920364830875769e-05
step: 230, loss: 0.002036026446148753
step: 240, loss: 0.0017653836403042078
step: 250, loss: 7.533238385803998e-05
step: 260, loss: 0.001752263866364956
step: 270, loss: 0.00157228484749794
step: 280, loss: 0.00018276485207024962
step: 290, loss: 0.0020487343426793814
step: 300, loss: 0.001448088907636702
step: 310, loss: 0.06412813812494278
step: 320, loss: 0.0003903310280293226
step: 330, loss: 0.0004977635107934475
step: 340, loss: 0.0015690050786361098
step: 350, loss: 0.0001572787296026945
step: 360, loss: 0.0005321606295183301
step: 370, loss: 0.002986299805343151
step: 380, loss: 0.0006917977589182556
step: 390, loss: 0.003634718945249915
step: 400, loss: 0.0004889382398687303
step: 410, loss: 0.011865470558404922
step: 420, loss: 0.005528326146304607
step: 430, loss: 0.0005814076284877956
step: 440, loss: 0.011765794828534126
step: 450, loss: 0.002238047309219837
step: 460, loss: 0.0060579534620046616
step: 470, loss: 0.004614208359271288
step: 480, loss: 0.004190150648355484
step: 490, loss: 0.00033790740417316556
step: 500, loss: 4.114386683795601e-05
step: 510, loss: 0.04483411833643913
step: 520, loss: 0.010486701503396034
step: 530, loss: 0.0009234509780071676
/usr/local/lib/python3.7/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=0.9503019043195542, f1=0.951851851851852,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 4.900001658825204e-05
step: 10, loss: 0.002125995932146907
step: 20, loss: 5.636244168272242e-05
step: 30, loss: 3.172205833834596e-05
step: 40, loss: 0.0005499008111655712
step: 50, loss: 0.0007578018121421337
step: 60, loss: 6.22076986473985e-05
step: 70, loss: 0.003953482024371624
step: 80, loss: 0.0006774947396479547
step: 90, loss: 8.697468001628295e-05
step: 100, loss: 0.0009831538191065192
step: 110, loss: 0.00010832230327650905
step: 120, loss: 3.612978252931498e-05
step: 130, loss: 5.762810178566724e-05
step: 140, loss: 7.358472066698596e-05
step: 150, loss: 0.12070897966623306
step: 160, loss: 6.587608368135989e-05
step: 170, loss: 0.001833026297390461
step: 180, loss: 0.0004802914336323738
step: 190, loss: 0.0009964696364477277
step: 200, loss: 2.2790532966610044e-05
step: 210, loss: 2.23129682126455e-05
step: 220, loss: 0.0045066229067742825
step: 230, loss: 0.0011522364802658558
step: 240, loss: 0.00024395008222199976
step: 250, loss: 0.00017486947763245553
step: 260, loss: 1.987011819437612e-05
step: 270, loss: 0.08380826562643051
step: 280, loss: 0.00019250721379648894
step: 290, loss: 9.97395472950302e-05
step: 300, loss: 1.7747035599313676e-05
step: 310, loss: 0.00010002597264247015
step: 320, loss: 0.012714577838778496
step: 330, loss: 0.0018542080651968718
step: 340, loss: 0.004367299377918243
step: 350, loss: 0.00037478230660781264
step: 360, loss: 0.07191342860460281
step: 370, loss: 0.01506004948168993
step: 380, loss: 0.0007565455161966383
step: 390, loss: 0.00031439459417015314
step: 400, loss: 9.290869638789445e-05
step: 410, loss: 0.0006506143836304545
step: 420, loss: 0.0015800727996975183
step: 430, loss: 0.0005568242631852627
step: 440, loss: 2.6020492441602983e-05
step: 450, loss: 0.0002402550890110433
```

```

step: 460, loss: 0.011134513653814793
step: 470, loss: 0.0047488734126091
step: 480, loss: 6.710709567414597e-05
step: 490, loss: 0.0035388455726206303
step: 500, loss: 0.0002701838093344122
step: 510, loss: 0.000900996383279562
step: 520, loss: 0.0006644055829383433
step: 530, loss: 0.00014153534721117467
/usr/local/lib/python3.7/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=0.9529953917050692, f1=0.9495412844036697,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005876995273865759
step: 10, loss: 0.0016890921397134662
step: 20, loss: 0.0015529242809861898
step: 30, loss: 0.0006882253219373524
step: 40, loss: 0.00013085087994113564
step: 50, loss: 0.017162904143333435
step: 60, loss: 0.00019212289771530777
step: 70, loss: 0.005926104262471199
step: 80, loss: 0.00030379253439605236
step: 90, loss: 0.00021660550555679947
step: 100, loss: 0.0003744593705050647
step: 110, loss: 0.0012202762300148606
step: 120, loss: 0.00013012989074923098
step: 130, loss: 0.0005564825842157006
step: 140, loss: 0.007954099215567112
step: 150, loss: 0.0019248765893280506
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.0010730619542300701
step: 170, loss: 0.00035237104748375714
step: 180, loss: 0.0002842945687007159
step: 190, loss: 0.0005066582816652954
step: 200, loss: 0.00011515174264786765
step: 210, loss: 0.0007228236645460129
step: 220, loss: 9.839794074650854e-05
step: 230, loss: 0.000205569202080369
step: 240, loss: 0.002754619810730219
step: 250, loss: 0.0018965171184390783
step: 260, loss: 0.0006380293052643538
step: 270, loss: 0.0016937876353040338
step: 280, loss: 0.0002230080426670611
step: 290, loss: 0.002621885621920228
step: 300, loss: 2.8918411771883257e-05

```

```

step: 310, loss: 0.005360602401196957
step: 320, loss: 0.00011724177602445707
step: 330, loss: 0.009377698414027691
step: 340, loss: 0.0009263679385185242
step: 350, loss: 0.0007410218240693212
step: 360, loss: 8.835447806632146e-05
step: 370, loss: 0.00036034415825270116
step: 380, loss: 0.0004560756788123399
step: 390, loss: 0.0007701104623265564
step: 400, loss: 0.0021432472858577967
step: 410, loss: 0.00019266194431111217
step: 420, loss: 0.00017850546282716095
step: 430, loss: 0.00146281230263412
step: 440, loss: 0.00857665203511715
step: 450, loss: 0.00012567559315357357
step: 460, loss: 0.02888459898531437
step: 470, loss: 3.195731551386416e-05
step: 480, loss: 0.00023267444339580834
step: 490, loss: 0.003944740165024996
step: 500, loss: 0.0009962029289454222
step: 510, loss: 0.04694445803761482
step: 520, loss: 0.0007015704759396613
step: 530, loss: 0.0021874357480555773
/usr/local/lib/python3.7/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=0.955113373438223, f1=0.9542302357836338,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001536218187538907
step: 10, loss: 0.0014512635534629226
step: 20, loss: 0.0013016954762861133
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.002939703641459346
step: 40, loss: 3.8341542676789686e-05
step: 50, loss: 0.00083001988241449
step: 60, loss: 0.0003462595632299781
step: 70, loss: 0.0001231642090715468
step: 80, loss: 0.00021546264179050922
step: 90, loss: 0.000424220081185922
step: 100, loss: 0.00010059009946417063
step: 110, loss: 0.00023865856928750873
step: 120, loss: 0.00015490267833229154
step: 130, loss: 0.00024830320035107434
step: 140, loss: 7.25828213035129e-05
step: 150, loss: 0.00018974195700138807

```



```

step: 160, loss: 0.0002033136843238026
step: 170, loss: 0.0012336552608758211
step: 180, loss: 0.00028843566542491317
step: 190, loss: 0.0017239681910723448
step: 200, loss: 0.0017183030722662807
step: 210, loss: 0.003830443602055311
step: 220, loss: 0.00017526136070955545
step: 230, loss: 0.017451802268624306
step: 240, loss: 0.0003075094136875123
step: 250, loss: 0.0009239893988706172
step: 260, loss: 0.00021902017761021852
step: 270, loss: 4.2647119698813185e-05
step: 280, loss: 0.00015021012222860008
step: 290, loss: 0.0001701346191111952
step: 300, loss: 0.0003480587329249829
step: 310, loss: 0.021543294191360474
step: 320, loss: 0.001782154431566596
step: 330, loss: 0.001007815240882337
step: 340, loss: 0.0003404179587960243
step: 350, loss: 0.0037923906929790974
step: 360, loss: 0.0003069978847634047
step: 370, loss: 0.0013777624117210507
step: 380, loss: 0.005356150213629007
step: 390, loss: 0.00010786540224216878
step: 400, loss: 0.0014728161040693521
step: 410, loss: 0.002208958612754941
step: 420, loss: 0.0005548345507122576
step: 430, loss: 2.5599441869417205e-05
step: 440, loss: 0.0002838941873051226
step: 450, loss: 0.02081458456814289
step: 460, loss: 0.0022130480501800776
step: 470, loss: 2.5617087885621004e-05
step: 480, loss: 0.0030296016484498978
step: 490, loss: 0.007963573560118675
step: 500, loss: 0.00204222253523767
step: 510, loss: 5.4422838729806244e-05
step: 520, loss: 2.4288239728775807e-05
step: 530, loss: 0.00015417239046655595
/usr/local/lib/python3.7/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=0.9534450651769087, f1=0.9534450651769087,
best_f1=0.9464867380176826
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:22, 253.58it/s]
load_f1 = 0.9552238805970149
real_f1 = 0.9553072625698324
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, 192.94it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

/usr/local/lib/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: 0.4441086947917938
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=0.2745098039215686, f1=0.2666666666666667,
best_f1=0.2666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4485799968242645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3636363636363636, f1=0.32558139534883723,
best_f1=0.32558139534883723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4765714108943939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37333333333333335, f1=0.33333333333333337,
best_f1=0.33333333333333337

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22310902178287506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5185185185185186, f1=0.43478260869565216,
best_f1=0.43478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23200663924217224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.6666666666666666, f1=0.5283018867924528,
best_f1=0.5283018867924528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17589926719665527
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7857142857142857, f1=0.7878787878787878,
best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32897961139678955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8888888888888889, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1437603235244751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9333333333333333, f1=0.7777777777777778,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0389041006565094
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.8484848484848484,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02432267554104328
/usr/local/lib/python3.7/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=0.9333333333333333, f1=0.8235294117647058,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11801612377166748
/usr/local/lib/python3.7/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=0.9333333333333333, f1=0.8235294117647058,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01933891512453556

```

```

/usr/local/lib/python3.7/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=0.8888888888888889, f1=0.9032258064516129,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036566019989550114
/usr/local/lib/python3.7/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=0.923076923076923, f1=0.896551724137931,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008069150149822235
/usr/local/lib/python3.7/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=0.923076923076923, f1=0.896551724137931,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015353811904788017
/usr/local/lib/python3.7/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=0.923076923076923, f1=0.896551724137931,
best_f1=0.7777777777777778
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, 109615.64it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.9285714285714286
real_f1 = 0.9285714285714286
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, 139.06it/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/

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

```
step: 0, loss: 0.5986150503158569
```

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: 0.437428742647171
```

```
step: 20, loss: 0.47875159978866577
```

```
step: 30, loss: 0.292853444814682
```

```
step: 40, loss: 0.3474624454975128
```

```
step: 50, loss: 0.6694056987762451
```

```
step: 60, loss: 0.4486309289932251
```

```
step: 70, loss: 0.5008935332298279
```

```
step: 80, loss: 0.49351394176483154
```

```
step: 90, loss: 0.1064935177564621
```

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

```
step: 100, loss: 0.29427096247673035
```

```
step: 110, loss: 0.03554417937994003
```

```

step: 120, loss: 0.09359300881624222
step: 130, loss: 0.04858642816543579
step: 140, loss: 0.1713612973690033
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 150, loss: 0.6492102146148682
step: 160, loss: 0.7175500392913818
step: 170, loss: 0.2375938594341278
step: 180, loss: 0.26093631982803345
step: 190, loss: 0.08939550071954727
step: 200, loss: 0.0518566332757473
step: 210, loss: 0.1956809014081955
step: 220, loss: 0.11975061148405075
step: 230, loss: 0.043445635586977005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9280898876404495, f1=0.9411764705882352,
best_f1=0.9411764705882352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02330072596669197
step: 10, loss: 0.23879721760749817
step: 20, loss: 0.03766654059290886
step: 30, loss: 0.025703255087137222
step: 40, loss: 0.04591492563486099
step: 50, loss: 0.006776393391191959
step: 60, loss: 0.019178269430994987
step: 70, loss: 0.006142381113022566
step: 80, loss: 0.007600728422403336
step: 90, loss: 0.027356429025530815
step: 100, loss: 0.0057134199887514114
step: 110, loss: 0.006934390868991613
step: 120, loss: 0.005452172365039587
step: 130, loss: 0.009504086337983608
step: 140, loss: 0.002959625329822302
step: 150, loss: 0.0905512124300003
step: 160, loss: 0.002129224594682455
step: 170, loss: 0.00428776303306222
step: 180, loss: 0.0024970516096800566
step: 190, loss: 0.009101220406591892
step: 200, loss: 0.006930846720933914
step: 210, loss: 0.032796453684568405
step: 220, loss: 0.00688403844833374

```

```

step: 230, loss: 0.0026395143941044807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9762174405436014, f1=0.9693530079455165,
best_f1=0.9693530079455165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012364122085273266
step: 10, loss: 0.004127959720790386
step: 20, loss: 0.12560059130191803
step: 30, loss: 0.011168353259563446
step: 40, loss: 0.023650314658880234
step: 50, loss: 0.020595744252204895
step: 60, loss: 0.032848965376615524
step: 70, loss: 0.0035041046794503927
step: 80, loss: 0.0035480724181979895
step: 90, loss: 0.0014240796444937587
step: 100, loss: 0.005538545548915863
step: 110, loss: 0.004837688058614731
step: 120, loss: 0.0018581593176349998
step: 130, loss: 0.0029818150214850903
step: 140, loss: 0.0012924716575071216
step: 150, loss: 0.005687990225851536
step: 160, loss: 0.003762314561754465
step: 170, loss: 0.0009528294904157519
step: 180, loss: 0.00925578735768795
step: 190, loss: 0.013990524224936962
step: 200, loss: 0.05940011888742447
step: 210, loss: 0.0020194929093122482
step: 220, loss: 0.005437001120299101
step: 230, loss: 0.008938489481806755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9855072463768116, f1=0.9887640449438202,
best_f1=0.9887640449438202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008377411402761936
step: 10, loss: 0.0019468506798148155
step: 20, loss: 0.021417254582047462
step: 30, loss: 0.001348039018921554
step: 40, loss: 0.06199260056018829
step: 50, loss: 0.06029415503144264
step: 60, loss: 0.005154056940227747
step: 70, loss: 0.04805527254939079
step: 80, loss: 0.0014830048894509673
step: 90, loss: 0.00966375321149826
step: 100, loss: 0.003974922467023134
step: 110, loss: 0.0015328233130276203
step: 120, loss: 0.006266390439122915
step: 130, loss: 0.0167236328125
step: 140, loss: 0.003015342867001891
step: 150, loss: 0.012900308705866337
step: 160, loss: 0.018568407744169235
step: 170, loss: 0.09484762698411942
step: 180, loss: 0.1901366114616394
step: 190, loss: 0.002857821062207222
step: 200, loss: 0.13463956117630005
step: 210, loss: 0.0022311853244900703
step: 220, loss: 0.0016971377190202475
step: 230, loss: 0.01570604369044304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987709497206704, f1=0.9842696629213483,
best_f1=0.9842696629213483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010767020285129547
step: 10, loss: 0.0021616709418594837
step: 20, loss: 0.007827876135706902
step: 30, loss: 0.0006198796327225864
step: 40, loss: 0.0011718962341547012
step: 50, loss: 0.0009238939383067191
step: 60, loss: 0.010563133284449577
step: 70, loss: 0.002111159497871995
step: 80, loss: 0.035512421280145645
step: 90, loss: 0.03665466606616974
step: 100, loss: 0.0013296942925080657

```

```

step: 110, loss: 0.002600919222459197
step: 120, loss: 0.0005536770913749933
step: 130, loss: 0.007429206743836403
step: 140, loss: 0.0008508638711646199
step: 150, loss: 0.10950541496276855
step: 160, loss: 0.0011781250359490514
step: 170, loss: 0.003503521205857396
step: 180, loss: 0.006399247795343399
step: 190, loss: 0.01619822531938553
step: 200, loss: 0.0014392693992704153
step: 210, loss: 0.009379712864756584
step: 220, loss: 0.0013142061652615666
step: 230, loss: 0.0016495289746671915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898305084745763, f1=0.9852774631936579,
best_f1=0.9852774631936579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005745846428908408
step: 10, loss: 0.0012881984002888203
step: 20, loss: 0.0019192147301509976
step: 30, loss: 0.0011815652251243591
step: 40, loss: 0.000644485407974571
step: 50, loss: 0.0018911109073087573
step: 60, loss: 0.008465372957289219
step: 70, loss: 0.041267555207014084
step: 80, loss: 0.011979545466601849
step: 90, loss: 0.03221826255321503
step: 100, loss: 0.010604233480989933
step: 110, loss: 0.05753196403384209
step: 120, loss: 0.0004539639921858907
step: 130, loss: 0.0005608039791695774
step: 140, loss: 0.00044666961184702814
step: 150, loss: 0.00019970069115515798
step: 160, loss: 0.0005320037598721683
step: 170, loss: 0.00037739056278951466
step: 180, loss: 0.0002497335080988705
step: 190, loss: 0.0003215629840269685
step: 200, loss: 0.0007763096946291625
step: 210, loss: 0.0007579156081192195
step: 220, loss: 0.00860938522964716

```

```

step: 230, loss: 0.002397526055574417
/usr/local/lib/python3.7/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=0.9887640449438202, f1=0.9831649831649831,
best_f1=0.9852774631936579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0043720100075006485
step: 10, loss: 0.001925653195939958
step: 20, loss: 0.0005878676893189549
step: 30, loss: 0.0004658487450797111
step: 40, loss: 0.00962341669946909
step: 50, loss: 0.0013213068014010787
step: 60, loss: 0.0009980887407436967
step: 70, loss: 0.0006793939392082393
step: 80, loss: 0.0005393655155785382
step: 90, loss: 0.007443943992257118
step: 100, loss: 0.00038619604310952127
step: 110, loss: 0.00043605134123936296
step: 120, loss: 0.0020200558938086033
step: 130, loss: 0.00048501763376407325
step: 140, loss: 0.0003466420457698405
step: 150, loss: 0.001715468824841082
step: 160, loss: 0.0003706439456436783
step: 170, loss: 0.001787995919585228
step: 180, loss: 0.0016773644601926208
step: 190, loss: 0.0009154635481536388
step: 200, loss: 0.004536330699920654
step: 210, loss: 0.0010872598504647613
step: 220, loss: 0.001035601831972599
step: 230, loss: 0.0011213443940505385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9899441340782122, f1=0.9866369710467707,
best_f1=0.9866369710467707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010940843494608998
step: 10, loss: 0.0031454931013286114
step: 20, loss: 0.002178513677790761

```

```

step: 30, loss: 0.001543754362501204
step: 40, loss: 0.0013773180544376373
step: 50, loss: 0.00089389638742432
step: 60, loss: 0.0009676739573478699
step: 70, loss: 0.00047599844401702285
step: 80, loss: 0.0025535474997013807
step: 90, loss: 0.0008063855348154902
step: 100, loss: 0.0008281863410957158
step: 110, loss: 0.005995786748826504
step: 120, loss: 0.0006905679474584758
step: 130, loss: 0.0016798492288216949
step: 140, loss: 0.0012139965547248721
step: 150, loss: 0.055972423404455185
step: 160, loss: 0.0010058985790237784
step: 170, loss: 0.025495145469903946
step: 180, loss: 0.000560527085326612
step: 190, loss: 0.0014482294209301472
step: 200, loss: 0.004643249325454235
step: 210, loss: 0.0021982192993164062
step: 220, loss: 0.0009736217907629907
step: 230, loss: 0.0006083755870349705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9910313901345291, f1=0.9899441340782122,
best_f1=0.9899441340782122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009659626521170139
step: 10, loss: 0.015620746649801731
step: 20, loss: 0.003918173722922802
step: 30, loss: 0.0033186727669090033
step: 40, loss: 0.002604302018880844
step: 50, loss: 0.0019232730846852064
step: 60, loss: 0.0006412536022253335
step: 70, loss: 0.04572899267077446
step: 80, loss: 0.0002674956922419369
step: 90, loss: 0.04350781440734863
step: 100, loss: 0.0012020451249554753
step: 110, loss: 0.00031153022428043187
step: 120, loss: 0.021587897092103958
step: 130, loss: 0.010849584825336933
step: 140, loss: 0.006792800035327673

```

```

step: 150, loss: 0.0025179157964885235
step: 160, loss: 0.0016005930956453085
step: 170, loss: 0.00046429800568148494
step: 180, loss: 0.0011975426459684968
step: 190, loss: 0.00031800440046936274
step: 200, loss: 0.0482562780380249
step: 210, loss: 0.009396489709615707
step: 220, loss: 0.0006280419765971601
step: 230, loss: 0.01820503920316696
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9818181818181818,
best_f1=0.9899441340782122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000781575683504343
step: 10, loss: 0.0011334308655932546
step: 20, loss: 0.008203436620533466
step: 30, loss: 0.001125470269471407
step: 40, loss: 0.0065939947962760925
step: 50, loss: 0.0010575600899755955
step: 60, loss: 0.0032194843515753746
step: 70, loss: 0.017780551686882973
step: 80, loss: 0.001046716351993382
step: 90, loss: 0.0003704601258505136
step: 100, loss: 0.0003488972724881023
step: 110, loss: 0.0011983182048425078
step: 120, loss: 0.0007494688616134226
step: 130, loss: 0.0010512638837099075
step: 140, loss: 0.0013313748640939593
step: 150, loss: 0.003981601446866989
step: 160, loss: 0.0002468646562192589
step: 170, loss: 0.0004390981630422175
step: 180, loss: 0.001309796585701406
step: 190, loss: 0.0004885682137683034
step: 200, loss: 0.0011013041948899627
step: 210, loss: 0.0008340232307091355
step: 220, loss: 0.0007838188903406262
step: 230, loss: 0.0003785596345551312
/usr/local/lib/python3.7/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=0.9899441340782122, f1=0.9854096520763187,
best_f1=0.9899441340782122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026122687268070877
step: 10, loss: 0.0005991901853121817
step: 20, loss: 0.0004501942894421518
step: 30, loss: 0.0003876909613609314
step: 40, loss: 0.00010953644232358783
step: 50, loss: 0.00043313196511007845
step: 60, loss: 0.008757821284234524
step: 70, loss: 0.0002516110544092953
step: 80, loss: 0.025491690263152122
step: 90, loss: 0.1997428685426712
step: 100, loss: 0.001432460150681436
step: 110, loss: 0.0014843029202893376
step: 120, loss: 0.0007134171319194138
step: 130, loss: 0.0006034146063029766
step: 140, loss: 0.005409674718976021
step: 150, loss: 0.0006965665379539132
step: 160, loss: 0.001219345023855567
step: 170, loss: 0.0009456115658394992
step: 180, loss: 0.001234028022736311
step: 190, loss: 0.0008496297523379326
step: 200, loss: 0.009834131225943565
step: 210, loss: 0.0013135402696207166
step: 220, loss: 0.0028814119286835194
step: 230, loss: 0.0004973949398845434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9899216125419933, f1=0.9853768278965129,
best_f1=0.9899441340782122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008213830878958106
step: 10, loss: 0.00047155481297522783
step: 20, loss: 0.01751573383808136
step: 30, loss: 0.02434460073709488
step: 40, loss: 0.0009721823153086007
step: 50, loss: 0.0012192653957754374
step: 60, loss: 0.0008998035918921232
step: 70, loss: 0.0006776357186026871
step: 80, loss: 0.00024527302593924105
step: 90, loss: 0.004185358993709087
step: 100, loss: 0.00042776763439178467
step: 110, loss: 0.0006356853991746902
step: 120, loss: 0.014551605097949505
step: 130, loss: 0.000442343094618991
step: 140, loss: 0.0029216918628662825

```

```

step: 150, loss: 0.0004557531210593879
step: 160, loss: 0.0011098040267825127
step: 170, loss: 0.0022940575145184994
step: 180, loss: 0.0002878221566788852
step: 190, loss: 0.019483210518956184
step: 200, loss: 0.0003792686911765486
step: 210, loss: 0.0013801919994875789
step: 220, loss: 0.010647124610841274
step: 230, loss: 0.00046116681187413633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9887892376681614, f1=0.9864864864864865,
best_f1=0.9899441340782122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14507590234279633
step: 10, loss: 0.0017968944739550352
step: 20, loss: 0.00264654029160738
step: 30, loss: 0.0006965854554437101
step: 40, loss: 0.001062873750925064
step: 50, loss: 0.0023483531549572945
step: 60, loss: 0.0005073578795418143
step: 70, loss: 0.0004209741309750825
step: 80, loss: 0.00042114281677640975
step: 90, loss: 0.0004596621438395232
step: 100, loss: 0.0008024718845263124
step: 110, loss: 0.0022877329029142857
step: 120, loss: 0.00045985504402779043
step: 130, loss: 0.0004803144547622651
step: 140, loss: 0.0007724910392425954
step: 150, loss: 0.00023157104442361742
step: 160, loss: 0.0010673730866983533
step: 170, loss: 0.0003926655917894095
step: 180, loss: 0.02211819961667061
step: 190, loss: 0.00038608809700235724
step: 200, loss: 0.00011728123354259878
step: 210, loss: 0.0005730480188503861
step: 220, loss: 0.00037020654417574406
step: 230, loss: 0.0004820296308025718
/usr/local/lib/python3.7/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=0.9910112359550561, f1=0.9864864864864865,
best_f1=0.9899441340782122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013405601494014263
step: 10, loss: 0.00021059571008663625
step: 20, loss: 0.00029618185362778604
step: 30, loss: 0.0004431159177329391
step: 40, loss: 0.00020809288253076375
step: 50, loss: 0.00018369172175880522
step: 60, loss: 0.00028426278731785715
step: 70, loss: 0.0002984091406688094
step: 80, loss: 0.0005739018088206649
step: 90, loss: 0.00043538701720535755
step: 100, loss: 0.0014456815551966429
step: 110, loss: 0.00036720771458931267
step: 120, loss: 9.456455882173032e-05
step: 130, loss: 0.0019315644167363644
step: 140, loss: 0.00037580178468488157
step: 150, loss: 0.00015624158550053835
step: 160, loss: 0.0003790505579672754
step: 170, loss: 0.0007815136923454702
step: 180, loss: 0.00025626292335800827
step: 190, loss: 0.00022459648607764393
step: 200, loss: 0.00031522411154583097
step: 210, loss: 0.00021452446526382118
step: 220, loss: 0.000308505492284894
step: 230, loss: 0.00023638817947357893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910514541387023, 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: 0.0031247520819306374
step: 10, loss: 0.0005855374038219452
step: 20, loss: 0.0006416055257432163
step: 30, loss: 0.000335197284584865
step: 40, loss: 0.00017385535466019064
step: 50, loss: 0.00023553898790851235
step: 60, loss: 0.040373463183641434
step: 70, loss: 0.00027634581783786416
step: 80, loss: 0.00023303908528760076
step: 90, loss: 0.0001714661775622517
step: 100, loss: 0.00015058100689202547

```

```

step: 110, loss: 0.0006340599502436817
step: 120, loss: 0.033273421227931976
step: 130, loss: 0.0001753996330080554
step: 140, loss: 0.010574061423540115
step: 150, loss: 0.0004936017212457955
step: 160, loss: 0.009251359850168228
step: 170, loss: 0.0001282918528886512
step: 180, loss: 0.0002731800777837634
step: 190, loss: 0.0001462164509575814
step: 200, loss: 0.0005214784760028124
step: 210, loss: 0.05918993800878525
step: 220, loss: 0.0002452232292853296
step: 230, loss: 0.0002824443799909204
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.9865470852017937,
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.")
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 [00:16, 150.24it/s]
load_f1 = 0.9910514541387023
real_f1 = 0.9910514541387023
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 134.61it/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/

```

```

Downloading: 100% 481/481 [00:00<00:00, 421kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 2.08MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 1.33MB/s]
Downloading: 100% 501M/501M [00:16<00:00, 29.8MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6288654804229736
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: 0.485752135515213
step: 20, loss: 0.28331759572029114
step: 30, loss: 0.3710800111293793
step: 40, loss: 0.3606804609298706
step: 50, loss: 0.46879276633262634
step: 60, loss: 0.2808845043182373
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.24882742762565613
step: 80, loss: 0.1061452180147171
step: 90, loss: 0.16169393062591553
step: 100, loss: 0.20055298507213593
step: 110, loss: 0.11784131079912186
step: 120, loss: 0.056401196867227554
step: 130, loss: 0.10306194424629211
step: 140, loss: 0.16840830445289612
step: 150, loss: 0.07738148421049118
step: 160, loss: 0.06213057041168213
step: 170, loss: 0.10593786835670471
step: 180, loss: 0.0631261020898819
step: 190, loss: 0.05551324039697647
step: 200, loss: 0.03768851235508919
step: 210, loss: 0.06582789123058319

```

```

step: 220, loss: 0.025098096579313278
step: 230, loss: 0.19528424739837646
step: 240, loss: 0.03357835114002228
step: 250, loss: 0.020905308425426483
step: 260, loss: 0.11049225926399231
step: 270, loss: 0.40649014711380005
step: 280, loss: 0.038020387291908264
step: 290, loss: 0.051344700157642365
step: 300, loss: 0.02441527508199215
step: 310, loss: 0.17161020636558533
step: 320, loss: 0.06892509758472443
step: 330, loss: 0.066946841776371
step: 340, loss: 0.4741300344467163
step: 350, loss: 0.05909913405776024
step: 360, loss: 0.035353418439626694
step: 370, loss: 0.13531748950481415
step: 380, loss: 0.07894548028707504
step: 390, loss: 0.01942877098917961
step: 400, loss: 0.04285210743546486
step: 410, loss: 0.25867530703544617
step: 420, loss: 0.03558626398444176
step: 430, loss: 0.022922836244106293
step: 440, loss: 0.0744064450263977
step: 450, loss: 0.025865288451313972
step: 460, loss: 0.026060443371534348
step: 470, loss: 0.02465568482875824
step: 480, loss: 0.1604047417640686
step: 490, loss: 0.05809386074542999
step: 500, loss: 0.011182593181729317
step: 510, loss: 0.020829351618885994
step: 520, loss: 0.07254165410995483
step: 530, loss: 0.01648685708642006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9521126760563381, f1=0.9475164011246484,
best_f1=0.9475164011246484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007375920191407204
step: 10, loss: 0.021509364247322083
step: 20, loss: 0.012565177865326405
step: 30, loss: 0.049572400748729706

```

step: 40, loss: 0.10987763106822968
step: 50, loss: 0.0639541894197464
step: 60, loss: 0.01274263858795166
step: 70, loss: 0.028926018625497818
step: 80, loss: 0.07182136178016663
step: 90, loss: 0.011865264736115932
step: 100, loss: 0.07077401876449585
step: 110, loss: 0.004545687232166529
step: 120, loss: 0.02102157287299633
step: 130, loss: 0.004590051714330912
step: 140, loss: 0.11584138125181198
step: 150, loss: 0.04499829187989235
step: 160, loss: 0.021218430250883102
step: 170, loss: 0.027092600241303444
step: 180, loss: 0.023820240050554276
step: 190, loss: 0.00984764751046896
step: 200, loss: 0.15885502099990845
step: 210, loss: 0.017788246273994446
step: 220, loss: 0.0008655391284264624
step: 230, loss: 0.04948872700333595
step: 240, loss: 0.010799573734402657
step: 250, loss: 0.08866888284683228
step: 260, loss: 0.015504537150263786
step: 270, loss: 0.010690338909626007
step: 280, loss: 0.044366318732500076
step: 290, loss: 0.042257554829120636
step: 300, loss: 0.035891976207494736
step: 310, loss: 0.022063428536057472
step: 320, loss: 0.11515361815690994
step: 330, loss: 0.03555217385292053
step: 340, loss: 0.14134590327739716
step: 350, loss: 0.0010994613403454423
step: 360, loss: 0.09644460678100586
step: 370, loss: 0.03727257624268532
step: 380, loss: 0.13346058130264282
step: 390, loss: 0.005984154529869556
step: 400, loss: 0.0022635995410382748
step: 410, loss: 0.004394629504531622
step: 420, loss: 0.04967455938458443
step: 430, loss: 0.2056221216917038
step: 440, loss: 0.06754137575626373
step: 450, loss: 0.0262723658233881
step: 460, loss: 0.019436625763773918
step: 470, loss: 0.009332889690995216
step: 480, loss: 0.005566806998103857
step: 490, loss: 0.03755531460046768
step: 500, loss: 0.0015074870316311717
step: 510, loss: 0.009149827994406223


```

step: 520, loss: 0.31164905428886414
step: 530, loss: 0.04809901490807533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9570093457943926, f1=0.9503280224929709,
best_f1=0.9503280224929709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1021430715918541
step: 10, loss: 0.06141393631696701
step: 20, loss: 0.0055670179426670074
step: 30, loss: 0.04320753365755081
step: 40, loss: 0.03085133619606495
step: 50, loss: 0.012348262593150139
step: 60, loss: 0.010684536769986153
step: 70, loss: 0.01187633816152811
step: 80, loss: 0.0356520339846611
step: 90, loss: 0.002936005825176835
step: 100, loss: 0.023342080414295197
step: 110, loss: 0.051795538514852524
step: 120, loss: 0.16097328066825867
step: 130, loss: 0.02863341011106968
step: 140, loss: 0.004346709232777357
step: 150, loss: 0.013132303953170776
step: 160, loss: 0.035943686962127686
step: 170, loss: 0.00846150703728199
step: 180, loss: 0.02155696228146553
step: 190, loss: 0.0022143416572362185
step: 200, loss: 0.021648457273840904
step: 210, loss: 0.02010447345674038
step: 220, loss: 0.037908896803855896
step: 230, loss: 0.017362268641591072
step: 240, loss: 0.01973164640367031
step: 250, loss: 0.0994701087474823
step: 260, loss: 0.1257481873035431
step: 270, loss: 0.001480588922277093
step: 280, loss: 0.14728112518787384
step: 290, loss: 0.03962913528084755
step: 300, loss: 0.05206221342086792
step: 310, loss: 0.04959098994731903
step: 320, loss: 0.13310766220092773
step: 330, loss: 0.0020625984761863947

```

```

step: 340, loss: 0.0070491679944098
step: 350, loss: 0.11771009862422943
step: 360, loss: 0.010557474568486214
step: 370, loss: 0.039899036288261414
step: 380, loss: 0.0025326248724013567
step: 390, loss: 0.009451109915971756
step: 400, loss: 0.011594494804739952
step: 410, loss: 0.14981108903884888
step: 420, loss: 0.008110428228974342
step: 430, loss: 0.03361588716506958
step: 440, loss: 0.23600099980831146
step: 450, loss: 0.030082523822784424
step: 460, loss: 0.00874000322073698
step: 470, loss: 0.022194111719727516
step: 480, loss: 0.04388481378555298
step: 490, loss: 0.029592303559184074
step: 500, loss: 0.09439331293106079
step: 510, loss: 0.057481423020362854
step: 520, loss: 0.004795457702130079
step: 530, loss: 0.0023743542842566967
/usr/local/lib/python3.7/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=0.9570041608876559, f1=0.9541454377026402,
best_f1=0.9503280224929709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008923492394387722
step: 10, loss: 0.005071285180747509
step: 20, loss: 0.09397538751363754
step: 30, loss: 0.004290894139558077
step: 40, loss: 0.007786482106894255
step: 50, loss: 0.034569401293992996
step: 60, loss: 0.0018956403946503997
step: 70, loss: 0.017949050292372704
step: 80, loss: 0.04766800254583359
step: 90, loss: 0.030985694378614426
step: 100, loss: 0.003020227188244462
step: 110, loss: 0.018508536741137505
step: 120, loss: 0.0012633237056434155
step: 130, loss: 0.017467347905039787
step: 140, loss: 0.09575466811656952
step: 150, loss: 0.0007847792003303766
step: 160, loss: 0.008460755459964275
step: 170, loss: 0.004979767370969057
step: 180, loss: 0.07661715149879456
step: 190, loss: 0.019649244844913483

```

```

step: 200, loss: 0.010711570270359516
step: 210, loss: 0.0011135389795526862
step: 220, loss: 0.0007359113660641015
step: 230, loss: 0.002977896947413683
step: 240, loss: 0.011273759417235851
step: 250, loss: 0.07792043685913086
step: 260, loss: 0.06866040080785751
step: 270, loss: 0.029013656079769135
step: 280, loss: 0.012775263749063015
step: 290, loss: 0.03432433307170868
step: 300, loss: 0.0007634527864865959
step: 310, loss: 0.0028983752708882093
step: 320, loss: 0.03945217654109001
step: 330, loss: 0.019871661439538002
step: 340, loss: 0.011653345078229904
step: 350, loss: 0.04750717803835869
step: 360, loss: 0.021839195862412453
step: 370, loss: 0.0025729620829224586
step: 380, loss: 0.0066999709233641624
step: 390, loss: 0.0011072495253756642
step: 400, loss: 0.014611970633268356
step: 410, loss: 0.009373494423925877
step: 420, loss: 0.035712096840143204
step: 430, loss: 0.0012473866809159517
step: 440, loss: 0.003122468013316393
step: 450, loss: 0.009600913152098656
step: 460, loss: 0.03440520912408829
step: 470, loss: 0.0022760394494980574
step: 480, loss: 0.013771703466773033
step: 490, loss: 0.004227389115840197
step: 500, loss: 0.012132317759096622
step: 510, loss: 0.013353602960705757
step: 520, loss: 0.08920002728700638
step: 530, loss: 0.1488313376903534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9619666048237476, f1=0.9583333333333334,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008619545958936214
step: 10, loss: 0.015982380136847496

```

step: 20, loss: 0.008676326833665371
step: 30, loss: 0.0065771141089499
step: 40, loss: 0.00035642884904518723
step: 50, loss: 0.06720943748950958
step: 60, loss: 0.0050038667395710945
step: 70, loss: 0.0011605401523411274
step: 80, loss: 0.00015351444017142057
step: 90, loss: 0.002078142948448658
step: 100, loss: 0.005282931495457888
step: 110, loss: 0.0070014093071222305
step: 120, loss: 0.009575250558555126
step: 130, loss: 0.0158398374915123
step: 140, loss: 0.02231290191411972
step: 150, loss: 0.0036264052614569664
step: 160, loss: 0.0013840912142768502
step: 170, loss: 0.0538913831114769
step: 180, loss: 0.00020489735470619053
step: 190, loss: 0.000326580397086218
step: 200, loss: 0.0656353011727333
step: 210, loss: 0.0006649476708844304
step: 220, loss: 0.0007289524073712528
step: 230, loss: 0.015232514590024948
step: 240, loss: 0.012021702714264393
step: 250, loss: 0.19332365691661835
step: 260, loss: 0.0021722731180489063
step: 270, loss: 0.0026950987521559
step: 280, loss: 0.004870534408837557
step: 290, loss: 0.006991798989474773
step: 300, loss: 0.0025212306063622236
step: 310, loss: 0.008832243271172047
step: 320, loss: 0.014347955584526062
step: 330, loss: 0.0010743682505562901
step: 340, loss: 0.001979903085157275
step: 350, loss: 0.00024263565137516707
step: 360, loss: 0.0001853038847912103
step: 370, loss: 0.0055735427886247635
step: 380, loss: 7.179443491622806e-05
step: 390, loss: 0.0034030864480882883
step: 400, loss: 0.019673945382237434
step: 410, loss: 0.0705779418349266
step: 420, loss: 0.13048741221427917
step: 430, loss: 0.018338993191719055
step: 440, loss: 0.0005381684168241918
step: 450, loss: 0.0058121755719184875
step: 460, loss: 0.0024176097940653563
step: 470, loss: 0.024399280548095703
step: 480, loss: 0.010414346121251583
step: 490, loss: 0.004355846904218197

```

step: 500, loss: 0.08070524036884308
step: 510, loss: 0.0001653901272220537
step: 520, loss: 0.009880658239126205
step: 530, loss: 0.10616761445999146
/usr/local/lib/python3.7/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=0.9531471840984383, f1=0.945127719962157,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013680912554264069
step: 10, loss: 0.00043116166489198804
step: 20, loss: 0.025926612317562103
step: 30, loss: 0.0006608190014958382
step: 40, loss: 0.0033731497824192047
step: 50, loss: 0.025479182600975037
step: 60, loss: 0.004995746072381735
step: 70, loss: 0.0014211790403351188
step: 80, loss: 0.0003195625904481858
step: 90, loss: 0.0012718064244836569
step: 100, loss: 0.0019857781007885933
step: 110, loss: 0.00138586584944278
step: 120, loss: 0.002225269563496113
step: 130, loss: 0.0004595747450366616
step: 140, loss: 0.0009750755853019655
step: 150, loss: 0.004969730041921139
step: 160, loss: 0.048224348574876785
step: 170, loss: 0.007383132353425026
step: 180, loss: 0.0018912100931629539
step: 190, loss: 0.013614681549370289
step: 200, loss: 0.001677177962847054
step: 210, loss: 0.009800484403967857
step: 220, loss: 0.0017821211367845535
step: 230, loss: 0.01675511710345745
step: 240, loss: 0.0005851208698004484
step: 250, loss: 0.020565304905176163
step: 260, loss: 0.00026173682999797165
step: 270, loss: 0.006181905511766672
step: 280, loss: 0.0008438306394964457
step: 290, loss: 0.0026773822028189898
step: 300, loss: 0.019606098532676697
step: 310, loss: 0.03567751124501228
step: 320, loss: 4.613134660758078e-05
step: 330, loss: 0.003009065752848983
step: 340, loss: 0.001607035519555211
step: 350, loss: 0.00041614138172008097

```

```

step: 360, loss: 0.03994009643793106
step: 370, loss: 0.058567240834236145
step: 380, loss: 0.0026286409702152014
step: 390, loss: 0.010806644335389137
step: 400, loss: 0.031233392655849457
step: 410, loss: 0.0002484846336301416
step: 420, loss: 0.007670980412513018
step: 430, loss: 0.00029558135429397225
step: 440, loss: 9.972853877115995e-05
step: 450, loss: 0.29839181900024414
step: 460, loss: 0.0018445905297994614
step: 470, loss: 0.0031095684971660376
step: 480, loss: 0.0029769924003630877
step: 490, loss: 0.006082551088184118
step: 500, loss: 0.00042393867624923587
step: 510, loss: 0.162167489528656
step: 520, loss: 0.0005199628067202866
step: 530, loss: 0.001841936493292451
/usr/local/lib/python3.7/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=0.9520225776105363, f1=0.9529137529137528,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008391509763896465
step: 10, loss: 0.003170083509758115
step: 20, loss: 0.0006649140268564224
step: 30, loss: 0.013729685917496681
step: 40, loss: 0.007522210944443941
step: 50, loss: 0.011040165089070797
step: 60, loss: 0.0005425261333584785
step: 70, loss: 0.00959438644349575
step: 80, loss: 0.004675949923694134
step: 90, loss: 0.0008914096979424357
step: 100, loss: 0.004719872027635574
step: 110, loss: 0.0031993950251489878
step: 120, loss: 0.0004374477721285075
step: 130, loss: 0.00035946242860518396
step: 140, loss: 0.0023253539111465216
step: 150, loss: 0.0002787890553008765
step: 160, loss: 0.0001190460316138342
step: 170, loss: 0.004686680622398853
step: 180, loss: 0.004621587228029966
step: 190, loss: 0.031061485409736633
step: 200, loss: 0.010214812122285366
step: 210, loss: 0.001265712664462626

```

```

step: 220, loss: 0.00025493730208836496
step: 230, loss: 0.00013336297706700861
step: 240, loss: 0.008098341524600983
step: 250, loss: 0.057149600237607956
step: 260, loss: 0.0025769411586225033
step: 270, loss: 0.001480455743148923
step: 280, loss: 0.001710417214781046
step: 290, loss: 0.0016426276415586472
step: 300, loss: 0.00024626089725643396
step: 310, loss: 0.0007276529213413596
step: 320, loss: 0.008714403957128525
step: 330, loss: 0.00010329752694815397
step: 340, loss: 0.005717779975384474
step: 350, loss: 0.0020238792058080435
step: 360, loss: 0.00040157814510166645
step: 370, loss: 0.009187649935483932
step: 380, loss: 0.00038644776213914156
step: 390, loss: 0.054463956505060196
step: 400, loss: 0.0018372363410890102
step: 410, loss: 0.00031586273689754307
step: 420, loss: 0.10101620107889175
step: 430, loss: 0.00018118097796104848
step: 440, loss: 0.002096605021506548
step: 450, loss: 0.00022585262195207179
step: 460, loss: 0.007837449200451374
step: 470, loss: 0.11613009124994278
step: 480, loss: 0.0035158745013177395
step: 490, loss: 0.004683271050453186
step: 500, loss: 0.002241687849164009
step: 510, loss: 0.00044814770808443427
step: 520, loss: 0.00018866370373871177
step: 530, loss: 0.0006141714402474463
/usr/local/lib/python3.7/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=0.9617537313432836, f1=0.9526462395543175,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005191736854612827
step: 10, loss: 0.02961893193423748
step: 20, loss: 0.0024846079759299755
step: 30, loss: 0.0010699874255806208
step: 40, loss: 9.189759293803945e-05
step: 50, loss: 0.0001544282422401011
step: 60, loss: 0.0006959975580684841
step: 70, loss: 0.00031467911321669817

```

step: 80, loss: 0.008678772486746311
step: 90, loss: 0.00016276637325063348
step: 100, loss: 0.00026162611902691424
step: 110, loss: 0.00031018556910566986
step: 120, loss: 0.0019675730727612972
step: 130, loss: 0.00035805144580081105
step: 140, loss: 0.09398727864027023
step: 150, loss: 0.00043326543527655303
step: 160, loss: 7.397210720228031e-05
step: 170, loss: 0.19236597418785095
step: 180, loss: 0.000547681120224297
step: 190, loss: 0.0019492190331220627
step: 200, loss: 0.003471170086413622
step: 210, loss: 0.04747158661484718
step: 220, loss: 0.0002050673501798883
step: 230, loss: 0.04638028144836426
step: 240, loss: 0.006002907175570726
step: 250, loss: 0.0004212288185954094
step: 260, loss: 7.018852193141356e-05
step: 270, loss: 0.00574386864900589
step: 280, loss: 0.0515776127576828
step: 290, loss: 0.006348894443362951
step: 300, loss: 2.6702062314143404e-05
step: 310, loss: 0.0011187158524990082
step: 320, loss: 0.00951178465038538
step: 330, loss: 0.0006394210504367948
step: 340, loss: 0.03455164656043053
step: 350, loss: 4.4453907321440056e-05
step: 360, loss: 0.01294828299432993
step: 370, loss: 0.15917760133743286
step: 380, loss: 4.802181865670718e-05
step: 390, loss: 0.002872125245630741
step: 400, loss: 0.0011683496413752437
step: 410, loss: 0.00024564500199630857
step: 420, loss: 0.0004167566657997668
step: 430, loss: 0.001516237505711615
step: 440, loss: 0.0007517560152336955
step: 450, loss: 0.0007123345858417451
step: 460, loss: 0.0009518884471617639
step: 470, loss: 0.08601110428571701
step: 480, loss: 0.04963746294379234
step: 490, loss: 0.0029669629875570536
step: 500, loss: 0.00012150818656664342
step: 510, loss: 0.0019791231025010347
step: 520, loss: 5.0692488002823666e-05
step: 530, loss: 8.947360765887424e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.


```
warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9589816124469591, f1=0.9520225776105363,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020772992866113782
step: 10, loss: 0.0054675983265042305
step: 20, loss: 0.0014760845806449652
step: 30, loss: 0.10992126166820526
step: 40, loss: 0.0017914437921717763
step: 50, loss: 4.0688646549824625e-05
step: 60, loss: 0.0007053015870042145
step: 70, loss: 6.263340765144676e-05
step: 80, loss: 0.0044413600116968155
step: 90, loss: 0.046301405876874924
step: 100, loss: 0.0005694974097423255
step: 110, loss: 0.14719367027282715
step: 120, loss: 0.002199801616370678
step: 130, loss: 0.0011227603536099195
step: 140, loss: 8.609126234659925e-05
step: 150, loss: 0.00045227553346194327
step: 160, loss: 0.0615769624710083
step: 170, loss: 0.015313155017793179
step: 180, loss: 0.0006557072629220784
step: 190, loss: 0.0001385386858601123
step: 200, loss: 9.078205039259046e-05
step: 210, loss: 0.000601985608227551
step: 220, loss: 0.0014698677696287632
step: 230, loss: 4.95625936309807e-05
step: 240, loss: 4.652983261621557e-05
step: 250, loss: 0.0001990606979234144
step: 260, loss: 0.00777369225397706
step: 270, loss: 0.0007169164018705487
step: 280, loss: 0.005959708243608475
step: 290, loss: 0.00013844427303411067
step: 300, loss: 0.0003978207241743803
step: 310, loss: 0.028467008844017982
step: 320, loss: 0.00023906584829092026
step: 330, loss: 0.0010050918208435178
step: 340, loss: 0.005105162505060434
step: 350, loss: 0.021212302148342133
step: 360, loss: 3.082812327193096e-05
step: 370, loss: 0.004435398615896702
step: 380, loss: 3.831901994999498e-05
step: 390, loss: 0.00017126963939517736
step: 400, loss: 0.0005504159489646554
step: 410, loss: 3.911848398274742e-05
```

```

step: 420, loss: 0.004496007692068815
step: 430, loss: 0.002580983331426978
step: 440, loss: 3.90659442928154e-05
step: 450, loss: 6.870651850476861e-05
step: 460, loss: 2.387491440458689e-05
step: 470, loss: 2.8106271201977506e-05
step: 480, loss: 0.0002666312793735415
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 490, loss: 0.0009421227150596678
step: 500, loss: 3.819788616965525e-05
step: 510, loss: 0.013328748755156994
step: 520, loss: 0.010153532028198242
step: 530, loss: 0.008513235487043858
/usr/local/lib/python3.7/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=0.9593647828117703, f1=0.9499072356215214,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014042336260899901
step: 10, loss: 0.002130882814526558
step: 20, loss: 0.00023825951211620122
step: 30, loss: 0.004904159344732761
step: 40, loss: 0.0006147826788946986
step: 50, loss: 0.0007795222918502986
step: 60, loss: 0.012836672365665436
step: 70, loss: 0.0001276818074984476
step: 80, loss: 0.00014772583381272852
step: 90, loss: 7.625278522027656e-05
step: 100, loss: 0.000627193134278059
step: 110, loss: 0.018966197967529297
step: 120, loss: 0.0008766999817453325
step: 130, loss: 0.00029674035613425076
step: 140, loss: 0.00012040563160553575
step: 150, loss: 0.002908754628151655
step: 160, loss: 0.07257287949323654
step: 170, loss: 0.000627129920758307
step: 180, loss: 0.001228217501193285
step: 190, loss: 7.516788900829852e-05
step: 200, loss: 0.009937817230820656
step: 210, loss: 0.009009354747831821
step: 220, loss: 0.0010433378629386425
step: 230, loss: 0.00011989203630946577
step: 240, loss: 0.00029205798637121916
step: 250, loss: 6.142845813883469e-05
step: 260, loss: 9.080642485059798e-05

```

```
step: 270, loss: 5.995026367600076e-05
step: 280, loss: 0.019666234031319618
step: 290, loss: 0.0004036401805933565
step: 300, loss: 0.0022753910161554813
step: 310, loss: 0.032699767500162125
step: 320, loss: 0.0014114441582933068
step: 330, loss: 0.00118592893704772
step: 340, loss: 4.0577440813649446e-05
step: 350, loss: 0.0010106436675414443
step: 360, loss: 3.0278446502052248e-05
step: 370, loss: 4.127691499888897e-05
step: 380, loss: 0.0015650197165086865
step: 390, loss: 0.013734833337366581
step: 400, loss: 9.096736903302372e-05
step: 410, loss: 3.324282806715928e-05
step: 420, loss: 2.6329586034989916e-05
step: 430, loss: 8.101965067908168e-05
step: 440, loss: 3.390553320059553e-05
step: 450, loss: 8.90507290023379e-05
step: 460, loss: 2.035473517025821e-05
step: 470, loss: 0.0075259460136294365
step: 480, loss: 7.393061969196424e-05
step: 490, loss: 4.9161786591866985e-05
step: 500, loss: 3.893276516464539e-05
step: 510, loss: 0.0003547094820532948
step: 520, loss: 2.2630712919635698e-05
step: 530, loss: 0.002078624675050378
/usr/local/lib/python3.7/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=0.9548872180451128, f1=0.9443920829406222,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.994627175212372e-05
step: 10, loss: 0.00033704566885717213
step: 20, loss: 0.0002637383877299726
step: 30, loss: 9.098292503040284e-05
step: 40, loss: 0.00010265718447044492
step: 50, loss: 8.394053293159232e-05
step: 60, loss: 4.238836481817998e-05
step: 70, loss: 4.938534402754158e-05
step: 80, loss: 0.0011116957757622004
step: 90, loss: 0.0013758731074631214
step: 100, loss: 0.0002265494258608669
step: 110, loss: 0.00036492408253252506
step: 120, loss: 0.006486534606665373
```

```
step: 130, loss: 7.097150955814868e-05
step: 140, loss: 9.942039469024166e-05
step: 150, loss: 0.00031834785477258265
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 5.248269735602662e-05
step: 170, loss: 0.0011363951489329338
step: 180, loss: 2.2925076336832717e-05
step: 190, loss: 0.007609707769006491
step: 200, loss: 0.00020381961076054722
step: 210, loss: 3.9736096368869767e-05
step: 220, loss: 0.05561896413564682
step: 230, loss: 6.113474955782294e-05
step: 240, loss: 0.00249930820427835
step: 250, loss: 2.751447209448088e-05
step: 260, loss: 0.002509907353669405
step: 270, loss: 0.00484563410282135
step: 280, loss: 0.008361740037798882
step: 290, loss: 4.9851521907839924e-05
step: 300, loss: 5.345950194168836e-05
step: 310, loss: 0.00016923079965636134
step: 320, loss: 0.0005543893203139305
step: 330, loss: 1.71063238667557e-05
step: 340, loss: 0.0020117871463298798
step: 350, loss: 0.0064307330176234245
step: 360, loss: 0.00014003601972945035
step: 370, loss: 0.002018512925133109
step: 380, loss: 0.0006623836234211922
step: 390, loss: 0.0012479767901822925
step: 400, loss: 3.62146929546725e-05
step: 410, loss: 0.0008088206523098052
step: 420, loss: 0.00011088775499956682
step: 430, loss: 0.0019038774771615863
step: 440, loss: 9.835469245444983e-05
step: 450, loss: 0.0007470593554899096
step: 460, loss: 0.00023644632892683148
step: 470, loss: 0.0005721452762372792
step: 480, loss: 0.0026767419185489416
step: 490, loss: 2.07454759220127e-05
step: 500, loss: 0.001840723678469658
step: 510, loss: 0.1617140918970108
step: 520, loss: 0.00044280808651819825
step: 530, loss: 0.07681327313184738
/usr/local/lib/python3.7/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=0.9556185080264401, f1=0.9472693032015067,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0005009963642805815
step: 10, loss: 0.006282491609454155
step: 20, loss: 0.0008013464394025505
step: 30, loss: 9.699313523015007e-05
step: 40, loss: 0.010990649461746216
step: 50, loss: 5.128288466949016e-05
step: 60, loss: 0.0004854490398429334
step: 70, loss: 0.0015692607266828418
step: 80, loss: 0.00010656502126948908
step: 90, loss: 0.0005560352001339197
step: 100, loss: 0.040564700961112976
step: 110, loss: 0.00036424002610147
step: 120, loss: 0.0006053808028809726
step: 130, loss: 0.0010282830335199833
step: 140, loss: 0.0002962018479593098
step: 150, loss: 0.00022675468062516302
step: 160, loss: 6.889387441333383e-05
step: 170, loss: 8.849461300997064e-05
step: 180, loss: 6.684961408609524e-05
step: 190, loss: 0.00037547049578279257
step: 200, loss: 0.001585376332513988
step: 210, loss: 2.9782717319903895e-05
step: 220, loss: 4.624500797945075e-05
step: 230, loss: 0.0015913958195596933
step: 240, loss: 0.000538148800842464
step: 250, loss: 9.697694622445852e-05
step: 260, loss: 0.0012368463212624192
step: 270, loss: 0.0019275617087259889
step: 280, loss: 9.23377665458247e-05
step: 290, loss: 0.0018358929082751274
step: 300, loss: 0.0006814718944951892
step: 310, loss: 0.00010303199815098196
step: 320, loss: 0.0002134514506906271
step: 330, loss: 0.0008830912411212921
step: 340, loss: 0.00012918405991513282
step: 350, loss: 0.0005657350993715227
step: 360, loss: 6.79103031870909e-05
step: 370, loss: 0.00731635931879282
step: 380, loss: 0.0012721432140097022
step: 390, loss: 0.00031537661561742425
step: 400, loss: 6.669623689958826e-05
step: 410, loss: 0.0017246264033019543
step: 420, loss: 0.0010294138919562101
step: 430, loss: 0.016807921230793
step: 440, loss: 9.25174099393189e-05
step: 450, loss: 0.002530076075345278
```

```

step: 460, loss: 5.416743806563318e-05
step: 470, loss: 0.0016481587663292885
step: 480, loss: 0.0027501245494931936
step: 490, loss: 2.6485391572350636e-05
step: 500, loss: 0.00011280539183644578
step: 510, loss: 0.03194297105073929
step: 520, loss: 0.00293164886534214
step: 530, loss: 3.959520108764991e-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=0.9549295774647887, f1=0.9502347417840376,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.4468517444329336e-05
step: 10, loss: 3.32954696204979e-05
step: 20, loss: 0.00010107879643328488
step: 30, loss: 1.4684928828501143e-05
step: 40, loss: 0.00027231633430346847
step: 50, loss: 0.0029896018095314503
step: 60, loss: 4.551037636701949e-05
step: 70, loss: 5.564386083278805e-05
step: 80, loss: 0.05303816869854927
step: 90, loss: 3.0694362067151815e-05
step: 100, loss: 4.2973792005795985e-05
step: 110, loss: 2.079747719108127e-05
step: 120, loss: 3.402197762625292e-05
step: 130, loss: 3.4222801332361996e-05
step: 140, loss: 5.200988380238414e-05
step: 150, loss: 0.0011727677192538977
step: 160, loss: 4.2526822653599083e-05
step: 170, loss: 0.0006791171617805958
step: 180, loss: 4.40297371824272e-05
step: 190, loss: 0.0007889374974183738
step: 200, loss: 0.00028823420871049166
step: 210, loss: 3.362025017850101e-05
step: 220, loss: 1.7564303561812267e-05
step: 230, loss: 0.000599399209022522
step: 240, loss: 0.00046852423110976815
step: 250, loss: 0.00012540147872641683
step: 260, loss: 1.7538308384246193e-05
step: 270, loss: 0.0001378945744363591
step: 280, loss: 1.5999728930182755e-05
step: 290, loss: 1.379817786073545e-05
step: 300, loss: 3.1112223950913176e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0

```

```

step: 310, loss: 0.07132457196712494
step: 320, loss: 1.43794213727233e-05
step: 330, loss: 2.9819282644893974e-05
step: 340, loss: 0.00014830502914264798
step: 350, loss: 1.0881412890739739e-05
step: 360, loss: 0.09512892365455627
step: 370, loss: 6.338894309010357e-05
step: 380, loss: 0.0004019229381810874
step: 390, loss: 0.0030987486243247986
step: 400, loss: 3.9595484849996865e-05
step: 410, loss: 1.8559087038738653e-05
step: 420, loss: 4.447386891115457e-05
step: 430, loss: 0.00010594706691335887
step: 440, loss: 1.599223105586134e-05
step: 450, loss: 3.5330151149537414e-05
step: 460, loss: 0.0032710740342736244
step: 470, loss: 0.0013077451149001718
step: 480, loss: 1.0289153578924015e-05
step: 490, loss: 0.00014513774658553302
step: 500, loss: 1.1011848073394503e-05
step: 510, loss: 2.4276361727970652e-05
step: 520, loss: 1.8067245036945678e-05
step: 530, loss: 1.1615328730840702e-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=0.9591078066914498, f1=0.9503480278422275,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009267584537155926
step: 10, loss: 0.0006069808732718229
step: 20, loss: 1.3462998140312266e-05
step: 30, loss: 0.00044861005153506994
step: 40, loss: 0.0008861593087203801
step: 50, loss: 3.755330908461474e-05
step: 60, loss: 1.8361513866693713e-05
step: 70, loss: 0.00010330192890251055
step: 80, loss: 0.0004495461762417108
step: 90, loss: 1.1697295121848583e-05
step: 100, loss: 1.6424557543359697e-05
step: 110, loss: 0.001236367505043745
step: 120, loss: 1.3328939530765638e-05
step: 130, loss: 0.00018047446792479604
step: 140, loss: 0.0002803850802592933
step: 150, loss: 0.0015726215206086636
step: 160, loss: 0.00015860980784054846

```

```

step: 170, loss: 0.008074691519141197
step: 180, loss: 4.913975135423243e-05
step: 190, loss: 1.8752602045424283e-05
step: 200, loss: 1.849194086389616e-05
step: 210, loss: 5.339948620530777e-05
step: 220, loss: 1.336989225819707e-05
step: 230, loss: 1.3947297702543437e-05
step: 240, loss: 0.0009717267821542919
step: 250, loss: 0.0011719915783032775
step: 260, loss: 1.8670574718271382e-05
step: 270, loss: 3.320625182823278e-05
step: 280, loss: 0.00040882715256884694
step: 290, loss: 1.1216681741643697e-05
step: 300, loss: 2.423105252091773e-05
step: 310, loss: 2.1728508727392182e-05
step: 320, loss: 2.0212455638102256e-05
step: 330, loss: 0.0010238600661978126
step: 340, loss: 0.0007582620601169765
step: 350, loss: 1.1533386896189768e-05
step: 360, loss: 0.0001450423151254654
step: 370, loss: 1.9292716388008557e-05
step: 380, loss: 0.0020846312399953604
step: 390, loss: 7.119657675502822e-05
step: 400, loss: 0.0013847503578290343
step: 410, loss: 1.6874499124241993e-05
step: 420, loss: 1.6644184142933227e-05
step: 430, loss: 4.43456374341622e-05
step: 440, loss: 1.71844512806274e-05
step: 450, loss: 1.383927337883506e-05
step: 460, loss: 0.030881071463227272
step: 470, loss: 1.3574805961980019e-05
step: 480, loss: 1.4118612853053492e-05
step: 490, loss: 1.391005571349524e-05
step: 500, loss: 0.0010308349737897515
step: 510, loss: 0.03429664671421051
step: 520, loss: 0.001221306505613029
step: 530, loss: 0.0006106395157985389
/usr/local/lib/python3.7/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=0.958139534883721, f1=0.9499072356215214,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.4871026905893814e-05
step: 10, loss: 0.0015870376955717802
step: 20, loss: 0.0005857566720806062

```


step: 30, loss: 0.020744040608406067
step: 40, loss: 0.0011731457198038697
step: 50, loss: 0.00509922718629241
step: 60, loss: 3.244240360800177e-05
step: 70, loss: 1.28185465655406e-05
step: 80, loss: 1.2557789887068793e-05
step: 90, loss: 1.8584767531137913e-05
step: 100, loss: 8.266358236141969e-06
step: 110, loss: 4.97374712722376e-05
step: 120, loss: 1.2751514987030532e-05
step: 130, loss: 1.2900490219180938e-05
step: 140, loss: 1.2621119822142646e-05
step: 150, loss: 1.2464678547985386e-05
step: 160, loss: 1.0568544894340448e-05
step: 170, loss: 1.389488170389086e-05
step: 180, loss: 3.183521039318293e-05
step: 190, loss: 0.0010871572885662317
step: 200, loss: 0.0012896961998194456
step: 210, loss: 1.326179062743904e-05
step: 220, loss: 1.1559453923837282e-05
step: 230, loss: 1.5232465557346586e-05
step: 240, loss: 1.1842569620057475e-05
step: 250, loss: 8.992789844342042e-06
step: 260, loss: 8.79535946296528e-06
step: 270, loss: 1.1883535989909433e-05
step: 280, loss: 1.4494728930003475e-05
step: 290, loss: 0.0020515206269919872
step: 300, loss: 9.294528354075737e-06
step: 310, loss: 0.019631147384643555
step: 320, loss: 1.5053633433126379e-05
step: 330, loss: 1.4271403415477835e-05
step: 340, loss: 1.3723773008678108e-05
step: 350, loss: 0.0027702045626938343
step: 360, loss: 0.00013896419841330498
step: 370, loss: 3.691388337756507e-05
step: 380, loss: 1.786195207387209e-05
step: 390, loss: 1.2710526789305732e-05
step: 400, loss: 0.006601858418434858
step: 410, loss: 2.1228694095043465e-05
step: 420, loss: 3.232613380532712e-05
step: 430, loss: 8.5569317889167e-06
step: 440, loss: 1.4706918591400608e-05
step: 450, loss: 0.002358623081818223
step: 460, loss: 0.0011143352603539824
step: 470, loss: 9.868217603070661e-06
step: 480, loss: 0.0023155766539275646
step: 490, loss: 0.000782908289693296
step: 500, loss: 0.000983199686743319

```

step: 510, loss: 9.64469654718414e-06
step: 520, loss: 1.2375253390928265e-05
step: 530, loss: 1.3280307030072436e-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=0.9592505854800936, f1=0.9510032664489033,
best_f1=0.9583333333333334
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 [00:25, 221.63it/s]
load_f1 = 0.9632044713553797
real_f1 = 0.9627906976744186
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 [00:25, 172.62it/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: 0.489595890045166
step: 10, loss: 0.4688612222671509
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.4299522936344147
step: 30, loss: 0.30164867639541626
step: 40, loss: 0.2963455617427826
step: 50, loss: 0.4180434048175812
step: 60, loss: 0.5037205219268799
step: 70, loss: 0.356368750333786
step: 80, loss: 0.34969931840896606
step: 90, loss: 0.2452327162027359
step: 100, loss: 0.2607819139957428
step: 110, loss: 0.28164908289909363
step: 120, loss: 0.43400999903678894
step: 130, loss: 0.16213463246822357
step: 140, loss: 0.26105380058288574
step: 150, loss: 0.2851618230342865
step: 160, loss: 0.27979281544685364
step: 170, loss: 0.14590629935264587
step: 180, loss: 0.23447485268115997
step: 190, loss: 0.25249674916267395
step: 200, loss: 0.11647598445415497
step: 210, loss: 0.28719401359558105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6144329896907217, f1=0.6479481641468683,
best_f1=0.6479481641468683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12456713616847992
step: 10, loss: 0.057955704629421234
step: 20, loss: 0.4066007733345032
step: 30, loss: 0.28574422001838684
step: 40, loss: 0.4026121199131012
step: 50, loss: 0.14538036286830902
step: 60, loss: 0.2171444445848465
step: 70, loss: 0.13043151795864105
step: 80, loss: 0.14937132596969604
step: 90, loss: 0.20859958231449127
step: 100, loss: 0.3205808401107788
step: 110, loss: 0.21256816387176514
step: 120, loss: 0.1851554661989212
step: 130, loss: 0.10339178144931793
step: 140, loss: 0.09407323598861694
step: 150, loss: 0.21792863309383392
step: 160, loss: 0.04971500486135483
step: 170, loss: 0.2866921126842499
step: 180, loss: 0.1544969230890274
step: 190, loss: 0.27149730920791626
step: 200, loss: 0.10041479021310806
step: 210, loss: 0.13467811048030853
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.674373795761079, f1=0.7454909819639278,
best_f1=0.7454909819639278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033213213086128235
step: 10, loss: 0.0138859236612916
step: 20, loss: 0.22557802498340607
step: 30, loss: 0.054209571331739426
step: 40, loss: 0.21998095512390137
step: 50, loss: 0.09349401295185089
step: 60, loss: 0.2950966954231262
step: 70, loss: 0.08949587494134903

```

```

step: 80, loss: 0.1364307850599289
step: 90, loss: 0.036243245005607605
step: 100, loss: 0.16782784461975098
step: 110, loss: 0.07725614309310913
step: 120, loss: 0.1692705750465393
step: 130, loss: 0.11739422380924225
step: 140, loss: 0.08528255671262741
step: 150, loss: 0.0917784795165062
step: 160, loss: 0.0673350989818573
step: 170, loss: 0.1832197606563568
step: 180, loss: 0.06963681429624557
step: 190, loss: 0.03599703684449196
step: 200, loss: 0.09956791996955872
step: 210, loss: 0.1259998232126236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.698744769874477, f1=0.7484662576687117,
best_f1=0.7484662576687117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016745444387197495
step: 10, loss: 0.0949116125702858
step: 20, loss: 0.0902349203824997
step: 30, loss: 0.08379363268613815
step: 40, loss: 0.01854020170867443
step: 50, loss: 0.1441315859556198
step: 60, loss: 0.06273625046014786
step: 70, loss: 0.05689762532711029
step: 80, loss: 0.02970222197473049
step: 90, loss: 0.10272897779941559
step: 100, loss: 0.17183366417884827
step: 110, loss: 0.45302054286003113
step: 120, loss: 0.159913569688797
step: 130, loss: 0.37731391191482544
step: 140, loss: 0.2568422257900238
step: 150, loss: 0.055149417370557785
step: 160, loss: 0.18993885815143585
step: 170, loss: 0.14028297364711761
step: 180, loss: 0.014731211587786674
step: 190, loss: 0.1067156121134758
step: 200, loss: 0.03836456313729286
step: 210, loss: 0.18405181169509888

```

```

/usr/local/lib/python3.7/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=0.667883211678832, f1=0.7343173431734317,
best_f1=0.7484662576687117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13521452248096466
step: 10, loss: 0.04460352659225464
step: 20, loss: 0.05022807791829109
step: 30, loss: 0.005478133447468281
step: 40, loss: 0.03520054370164871
step: 50, loss: 0.11086012423038483
step: 60, loss: 0.08007320761680603
step: 70, loss: 0.14343883097171783
step: 80, loss: 0.09147017449140549
step: 90, loss: 0.03968989476561546
step: 100, loss: 0.05347424000501633
step: 110, loss: 0.051927078515291214
step: 120, loss: 0.06276608258485794
step: 130, loss: 0.012839713133871555
step: 140, loss: 0.12647587060928345
step: 150, loss: 0.042511288076639175
step: 160, loss: 0.02042296528816223
step: 170, loss: 0.05683651566505432
step: 180, loss: 0.016638778150081635
step: 190, loss: 0.10188252478837967
step: 200, loss: 0.05197429284453392
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 210, loss: 0.0940566211938858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7326732673267327, f1=0.73046875, best_f1=0.73046875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006449021864682436
step: 10, loss: 0.06466566026210785
step: 20, loss: 0.02220088429749012
step: 30, loss: 0.06193147599697113
step: 40, loss: 0.022060168907046318
step: 50, loss: 0.036712877452373505

```

```

step: 60, loss: 0.11024972051382065
step: 70, loss: 0.05372944846749306
step: 80, loss: 0.04795607179403305
step: 90, loss: 0.027353225275874138
step: 100, loss: 0.10432229936122894
step: 110, loss: 0.10500200092792511
step: 120, loss: 0.0068630650639534
step: 130, loss: 0.022127723321318626
step: 140, loss: 0.09093409031629562
step: 150, loss: 0.009828167036175728
step: 160, loss: 0.12771232426166534
step: 170, loss: 0.0481189526617527
step: 180, loss: 0.03214828670024872
step: 190, loss: 0.06657685339450836
step: 200, loss: 0.0822693407535553
step: 210, loss: 0.04545190930366516
/usr/local/lib/python3.7/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=0.72265625, f1=0.7265625, best_f1=0.73046875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07219663262367249
step: 10, loss: 0.006248844787478447
step: 20, loss: 0.009404178708791733
step: 30, loss: 0.01806035451591015
step: 40, loss: 0.013355453498661518
step: 50, loss: 0.1197333112359047
step: 60, loss: 0.02387571521103382
step: 70, loss: 0.03083282709121704
step: 80, loss: 0.02841382846236229
step: 90, loss: 0.03998001292347908
step: 100, loss: 0.06724371016025543
step: 110, loss: 0.06296724826097488
step: 120, loss: 0.0865708664059639
step: 130, loss: 0.034815434366464615
step: 140, loss: 0.027271540835499763
step: 150, loss: 0.040517471730709076
step: 160, loss: 0.19400519132614136
step: 170, loss: 0.01618693396449089
step: 180, loss: 0.01341017335653305
step: 190, loss: 0.015939651057124138
step: 200, loss: 0.05366664379835129
step: 210, loss: 0.06676206737756729
/usr/local/lib/python3.7/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=0.7237354085603112, f1=0.7480916030534351, best_f1=0.73046875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04354578256607056
step: 10, loss: 0.01512990240007639
step: 20, loss: 0.04782100021839142
step: 30, loss: 0.06900456547737122
step: 40, loss: 0.010114345699548721
step: 50, loss: 0.00018546140927355736
step: 60, loss: 0.00481441942974925
step: 70, loss: 0.13331779837608337
step: 80, loss: 0.021271884441375732
step: 90, loss: 0.07240837067365646
step: 100, loss: 0.1415664255619049
step: 110, loss: 0.007988505065441132
step: 120, loss: 0.09494996070861816
step: 130, loss: 0.008860225789248943
step: 140, loss: 0.07714192569255829
step: 150, loss: 0.07112086564302444
step: 160, loss: 0.126455619931221
step: 170, loss: 0.3235185742378235
step: 180, loss: 0.01537083275616169
step: 190, loss: 0.09147471189498901
step: 200, loss: 0.056962285190820694
step: 210, loss: 0.041189029812812805
/usr/local/lib/python3.7/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=0.7196969696969697, f1=0.7210626185958253, best_f1=0.73046875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007119270507246256
step: 10, loss: 0.127573162317276
step: 20, loss: 0.08784400671720505
step: 30, loss: 0.022670084610581398
step: 40, loss: 0.00441959360614419
step: 50, loss: 0.012868404388427734
step: 60, loss: 0.0016765424516052008
step: 70, loss: 0.07480868697166443
step: 80, loss: 0.0024582736659795046
step: 90, loss: 0.008170456625521183
step: 100, loss: 0.02023073472082615
step: 110, loss: 0.05631955340504646
step: 120, loss: 0.1782134771347046
step: 130, loss: 0.04616710916161537
step: 140, loss: 0.06995116919279099

```

```

step: 150, loss: 0.005863097496330738
step: 160, loss: 0.12710323929786682
step: 170, loss: 0.03719757869839668
step: 180, loss: 0.03188275173306465
step: 190, loss: 0.02284972183406353
step: 200, loss: 0.000769426638726145
step: 210, loss: 0.0034222430549561977
/usr/local/lib/python3.7/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=0.725233644859813, f1=0.714548802946593, best_f1=0.73046875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037877201102674007
step: 10, loss: 0.0011405437253415585
step: 20, loss: 0.009662244468927383
step: 30, loss: 0.019324814900755882
step: 40, loss: 0.0075689381919801235
step: 50, loss: 0.0015137175796553493
step: 60, loss: 0.0004933548625558615
step: 70, loss: 0.016874320805072784
step: 80, loss: 0.002773116109892726
step: 90, loss: 0.11677293479442596
step: 100, loss: 0.013860554434359074
step: 110, loss: 0.0007896493189036846
step: 120, loss: 0.004217525478452444
step: 130, loss: 0.008570320904254913
step: 140, loss: 0.012038270942866802
step: 150, loss: 0.020404228940606117
step: 160, loss: 0.014659654349088669
step: 170, loss: 0.08686266839504242
step: 180, loss: 0.0558755025267601
step: 190, loss: 0.012614517472684383
step: 200, loss: 0.02927352301776409
step: 210, loss: 0.014909583143889904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7414829659318636, f1=0.7489878542510122,
best_f1=0.7489878542510122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.05294830724596977
step: 10, loss: 0.0018503974424675107
step: 20, loss: 0.005562044680118561
step: 30, loss: 0.07092233747243881
step: 40, loss: 0.008914847858250141
step: 50, loss: 0.002594292862340808
step: 60, loss: 0.029122307896614075
step: 70, loss: 0.06055229902267456
step: 80, loss: 0.05980440601706505
step: 90, loss: 0.1320277899503708
step: 100, loss: 0.016573438420891762
step: 110, loss: 0.06185906007885933
step: 120, loss: 0.004811078310012817
step: 130, loss: 0.001345002674497664
step: 140, loss: 0.11359360814094543
step: 150, loss: 0.020076654851436615
step: 160, loss: 0.0005342342192307115
step: 170, loss: 0.019863640889525414
step: 180, loss: 0.006576809100806713
step: 190, loss: 0.04422573745250702
step: 200, loss: 0.00471603125333786
step: 210, loss: 0.007284762803465128
/usr/local/lib/python3.7/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=0.7215686274509804, f1=0.7389558232931727,
best_f1=0.7489878542510122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01454472728073597
step: 10, loss: 0.015679193660616875
step: 20, loss: 0.03119833767414093
step: 30, loss: 0.015594062395393848
step: 40, loss: 0.0005881953402422369
step: 50, loss: 0.0017578091938048601
step: 60, loss: 7.555213960586116e-05
step: 70, loss: 0.004979401361197233
step: 80, loss: 0.14965319633483887
step: 90, loss: 0.0020251499954611063
step: 100, loss: 7.053139415802434e-05
step: 110, loss: 0.008784355595707893
step: 120, loss: 0.0009313209448009729
step: 130, loss: 0.002837198553606868
step: 140, loss: 0.0014258098090067506
step: 150, loss: 6.496254354715347e-05
step: 160, loss: 0.0073502808809280396
step: 170, loss: 0.06938077509403229

```

```

step: 180, loss: 0.05249228700995445
step: 190, loss: 0.000216770960832946
step: 200, loss: 0.0026306959334760904
step: 210, loss: 0.0012155025033280253
/usr/local/lib/python3.7/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=0.7325581395348837, f1=0.72552783109405,
best_f1=0.7489878542510122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005543014500290155
step: 10, loss: 0.011137532070279121
step: 20, loss: 0.00020468181173782796
step: 30, loss: 0.0003237736818846315
step: 40, loss: 0.001957204192876816
step: 50, loss: 0.06704355776309967
step: 60, loss: 0.0020403279922902584
step: 70, loss: 0.0003207492991350591
step: 80, loss: 0.08206717669963837
step: 90, loss: 0.024105805903673172
step: 100, loss: 0.00034802238224074244
step: 110, loss: 0.01956026256084442
step: 120, loss: 0.007550851441919804
step: 130, loss: 0.0007393896230496466
step: 140, loss: 0.0276772603392601
step: 150, loss: 0.039623405784368515
step: 160, loss: 0.001270483830012381
step: 170, loss: 0.0010057003237307072
step: 180, loss: 0.0002839818480424583
step: 190, loss: 0.02134580723941326
step: 200, loss: 0.024930503219366074
step: 210, loss: 0.01382800005376339
/usr/local/lib/python3.7/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=0.7243460764587525, f1=0.7401574803149608,
best_f1=0.7489878542510122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000439988449215889
step: 10, loss: 0.040472518652677536
step: 20, loss: 0.0003227615379728377
step: 30, loss: 0.00035999788087792695
step: 40, loss: 0.015081135556101799
step: 50, loss: 0.0004436106246430427

```

```

step: 60, loss: 0.20361551642417908
step: 70, loss: 0.0013200975954532623
step: 80, loss: 0.04433615133166313
step: 90, loss: 9.630506974644959e-05
step: 100, loss: 0.0006564825889654458
step: 110, loss: 0.005210293922573328
step: 120, loss: 6.841322465334088e-05
step: 130, loss: 0.014993997290730476
step: 140, loss: 0.00010641217522788793
step: 150, loss: 0.001253564958460629
step: 160, loss: 6.873231177451089e-05
step: 170, loss: 0.09267804771661758
step: 180, loss: 0.0001136544524342753
step: 190, loss: 0.06481095403432846
step: 200, loss: 0.0006586111849173903
step: 210, loss: 0.0027367589063942432
/usr/local/lib/python3.7/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=0.7405940594059407, f1=0.7322834645669292,
best_f1=0.7489878542510122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002215983346104622
step: 10, loss: 0.0002641493920236826
step: 20, loss: 0.011582654900848866
step: 30, loss: 0.0003516385331749916
step: 40, loss: 0.007279830984771252
step: 50, loss: 7.188069866970181e-05
step: 60, loss: 0.048898059874773026
step: 70, loss: 0.005453081335872412
step: 80, loss: 0.0007416590815410018
step: 90, loss: 0.010442806407809258
step: 100, loss: 0.0015775029314681888
step: 110, loss: 0.011578788980841637
step: 120, loss: 0.0003819370467681438
step: 130, loss: 0.007175360340625048
step: 140, loss: 0.022584959864616394
step: 150, loss: 0.0024439177941530943
step: 160, loss: 0.010399842634797096
step: 170, loss: 0.02249552309513092
step: 180, loss: 0.0003195013850927353
step: 190, loss: 0.0056020962074398994
step: 200, loss: 0.0002986967738252133
step: 210, loss: 0.03916905075311661
/usr/local/lib/python3.7/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=0.7337278106508875, f1=0.7351778656126481,
best_f1=0.7489878542510122
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.")
2293it [00:06, 362.83it/s]
load_f1 = 0.7419354838709677
real_f1 = 0.7309236947791166
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, 170.32it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

```

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

```

```

# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Walmart-Amazon \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm 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/

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.5017642974853516
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.39835184812545776
step: 20, loss: 0.32629314064979553
step: 30, loss: 0.36360079050064087
step: 40, loss: 0.2780287563800812
step: 50, loss: 0.3146347999572754
step: 60, loss: 0.45127204060554504
step: 70, loss: 0.45228999853134155
step: 80, loss: 0.16489827632904053
step: 90, loss: 0.31304246187210083
step: 100, loss: 0.49461817741394043
step: 110, loss: 0.239065021276474
step: 120, loss: 0.3465840816497803
step: 130, loss: 0.30708053708076477
step: 140, loss: 0.14413447678089142
step: 150, loss: 0.3100510239601135
step: 160, loss: 0.20329086482524872
step: 170, loss: 0.3613343834877014
step: 180, loss: 0.17125721275806427
step: 190, loss: 0.19570714235305786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.39081084728240967

```



```

step: 10, loss: 0.3004913926124573
step: 20, loss: 0.5835643410682678
step: 30, loss: 0.23990567028522491
step: 40, loss: 0.556373119354248
step: 50, loss: 0.30839377641677856
step: 60, loss: 0.41895243525505066
step: 70, loss: 0.324418306350708
step: 80, loss: 0.15637117624282837
step: 90, loss: 0.2774491608142853
step: 100, loss: 0.18207064270973206
step: 110, loss: 0.3136231005191803
step: 120, loss: 0.20758749544620514
step: 130, loss: 0.41665178537368774
step: 140, loss: 0.25979024171829224
step: 150, loss: 0.2664256691932678
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.23249658942222595
step: 170, loss: 0.2905762195587158
step: 180, loss: 0.14131805300712585
step: 190, loss: 0.20303310453891754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5836065573770491, f1=0.6095238095238096,
best_f1=0.6095238095238096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26497676968574524
step: 10, loss: 0.2735713720321655
step: 20, loss: 0.6874564290046692
step: 30, loss: 0.16019099950790405
step: 40, loss: 0.1112205907702446
step: 50, loss: 0.20637401938438416
step: 60, loss: 0.04597248136997223
step: 70, loss: 0.29678189754486084
step: 80, loss: 0.17111310362815857
step: 90, loss: 0.1035386249423027
step: 100, loss: 0.13438235223293304
step: 110, loss: 0.4409577548503876
step: 120, loss: 0.34299561381340027
step: 130, loss: 0.14174094796180725
step: 140, loss: 0.07500474154949188
step: 150, loss: 0.24135909974575043

```

```

step: 160, loss: 0.20716947317123413
step: 170, loss: 0.16553093492984772
step: 180, loss: 0.21916906535625458
step: 190, loss: 0.22161976993083954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7073791348600508, f1=0.7291139240506329,
best_f1=0.7291139240506329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09052254259586334
step: 10, loss: 0.10102132707834244
step: 20, loss: 0.27280548214912415
step: 30, loss: 0.07709045708179474
step: 40, loss: 0.1994725614786148
step: 50, loss: 0.03955293074250221
step: 60, loss: 0.11476422101259232
step: 70, loss: 0.04205550253391266
step: 80, loss: 0.0562259815633297
step: 90, loss: 0.03859611600637436
step: 100, loss: 0.20566385984420776
step: 110, loss: 0.1846601665019989
step: 120, loss: 0.0735015869140625
step: 130, loss: 0.10251017659902573
step: 140, loss: 0.27571335434913635
step: 150, loss: 0.04946431145071983
step: 160, loss: 0.1443895548582077
step: 170, loss: 0.10362958163022995
step: 180, loss: 0.0607362799346447
step: 190, loss: 0.10914219915866852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7971014492753624, f1=0.79236276849642,
best_f1=0.79236276849642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.06532828509807587
step: 10, loss: 0.1093771681189537
step: 20, loss: 0.03185059130191803
step: 30, loss: 0.027860863134264946
step: 40, loss: 0.040596287697553635
step: 50, loss: 0.08037417382001877
step: 60, loss: 0.1198253408074379
step: 70, loss: 0.30532142519950867
step: 80, loss: 0.0348205491900444
step: 90, loss: 0.12378790974617004
step: 100, loss: 0.17026841640472412
step: 110, loss: 0.10880277305841446
step: 120, loss: 0.08680876344442368
step: 130, loss: 0.302944540977478
step: 140, loss: 0.06139629706740379
step: 150, loss: 0.20420779287815094
step: 160, loss: 0.10530388355255127
step: 170, loss: 0.07768961042165756
step: 180, loss: 0.04330798238515854
step: 190, loss: 0.02718883939087391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8157248157248156, f1=0.8126520681265206,
best_f1=0.8126520681265206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24034619331359863
step: 10, loss: 0.041461579501628876
step: 20, loss: 0.0021754775661975145
step: 30, loss: 0.041404008865356445
step: 40, loss: 0.00891118310391903
step: 50, loss: 0.09486515820026398
step: 60, loss: 0.08213811367750168
step: 70, loss: 0.024498235434293747
step: 80, loss: 0.19753704965114594
step: 90, loss: 0.13682982325553894
step: 100, loss: 0.14392134547233582
step: 110, loss: 0.058636441826820374
step: 120, loss: 0.03751114383339882
step: 130, loss: 0.1676187962293625
step: 140, loss: 0.09016726911067963
step: 150, loss: 0.008549720048904419

```

```

step: 160, loss: 0.09507094323635101
step: 170, loss: 0.26311194896698
step: 180, loss: 0.10081314295530319
step: 190, loss: 0.03937624394893646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8415841584158414, f1=0.8039215686274511,
best_f1=0.8039215686274511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12101537734270096
step: 10, loss: 0.07422280311584473
step: 20, loss: 0.003083351533859968
step: 30, loss: 0.007286530919373035
step: 40, loss: 0.008820497430860996
step: 50, loss: 0.12178301811218262
step: 60, loss: 0.17198151350021362
step: 70, loss: 0.005817511584609747
step: 80, loss: 0.0069188326597213745
step: 90, loss: 0.004423535894602537
step: 100, loss: 0.2558431029319763
step: 110, loss: 0.1649453192949295
step: 120, loss: 0.014516517519950867
step: 130, loss: 0.004290441051125526
step: 140, loss: 0.11207704991102219
step: 150, loss: 0.028589027002453804
step: 160, loss: 0.060946255922317505
step: 170, loss: 0.10118041932582855
step: 180, loss: 0.029926437884569168
step: 190, loss: 0.1679106056690216
/usr/local/lib/python3.7/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=0.8316326530612245, f1=0.8256410256410257,
best_f1=0.8039215686274511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0671127662062645
step: 10, loss: 0.028623955324292183
step: 20, loss: 0.004368809051811695
step: 30, loss: 0.13632997870445251

```

```

step: 40, loss: 0.172194242477417
step: 50, loss: 0.012117399834096432
step: 60, loss: 0.04913352429866791
step: 70, loss: 0.01950293406844139
step: 80, loss: 0.14511464536190033
step: 90, loss: 0.03770885989069939
step: 100, loss: 0.05274484306573868
step: 110, loss: 0.034986771643161774
step: 120, loss: 0.003709899028763175
step: 130, loss: 0.005618653725832701
step: 140, loss: 0.006136788055300713
step: 150, loss: 0.02327822521328926
step: 160, loss: 0.0031283816788345575
step: 170, loss: 0.001707402872852981
step: 180, loss: 0.1300620287656784
step: 190, loss: 0.007783690933138132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8475452196382429, f1=0.8390501319261214,
best_f1=0.8390501319261214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01467854157090187
step: 10, loss: 0.0008874769555404782
step: 20, loss: 0.005808142013847828
step: 30, loss: 0.016743525862693787
step: 40, loss: 0.30992239713668823
step: 50, loss: 0.011373069137334824
step: 60, loss: 0.006605310831218958
step: 70, loss: 0.12716367840766907
step: 80, loss: 0.1716408133506775
step: 90, loss: 0.02137431688606739
step: 100, loss: 0.005096845794469118
step: 110, loss: 0.0037259708624333143
step: 120, loss: 0.1340133100748062
step: 130, loss: 0.003939711023122072
step: 140, loss: 0.007404468487948179
step: 150, loss: 0.00418108468875289
step: 160, loss: 0.0026083479169756174
step: 170, loss: 0.00856892578303814
step: 180, loss: 0.022337552160024643
step: 190, loss: 0.002105828607454896

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8511749347258485, f1=0.8351063829787233,
best_f1=0.8351063829787233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002034214325249195
step: 10, loss: 0.003281916258856654
step: 20, loss: 0.04076502472162247
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 30, loss: 0.29277798533439636
step: 40, loss: 0.002943199360743165
step: 50, loss: 0.006343236193060875
step: 60, loss: 0.006046496797353029
step: 70, loss: 0.002578162122517824
step: 80, loss: 0.0034300487022846937
step: 90, loss: 0.002149517647922039
step: 100, loss: 0.02926124632358551
step: 110, loss: 0.061554595828056335
step: 120, loss: 0.0035881372168660164
step: 130, loss: 0.02835085801780224
step: 140, loss: 0.0014636663254350424
step: 150, loss: 0.0011886701686307788
step: 160, loss: 0.006882074289023876
step: 170, loss: 0.001932156621478498
step: 180, loss: 0.016427261754870415
step: 190, loss: 0.017361978068947792
/usr/local/lib/python3.7/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=0.8333333333333334, f1=0.8368421052631578,
best_f1=0.8351063829787233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004159674048423767
step: 10, loss: 0.0005387195269577205
step: 20, loss: 0.0007424574578180909
step: 30, loss: 0.0967484712600708
step: 40, loss: 0.017843348905444145
step: 50, loss: 0.03052736446261406
step: 60, loss: 0.005166177172213793

```

```

step: 70, loss: 0.01655563898384571
step: 80, loss: 0.011969886720180511
step: 90, loss: 0.0013037441531196237
step: 100, loss: 0.00041636289097368717
step: 110, loss: 0.23640434443950653
step: 120, loss: 0.019326679408550262
step: 130, loss: 0.0018370550824329257
step: 140, loss: 0.08699968457221985
step: 150, loss: 0.0026995448861271143
step: 160, loss: 0.0009922342142090201
step: 170, loss: 0.0027926110196858644
step: 180, loss: 0.011871028691530228
step: 190, loss: 0.0005155610851943493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8548812664907651, f1=0.851063829787234,
best_f1=0.851063829787234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014581084251403809
step: 10, loss: 0.005432508885860443
step: 20, loss: 0.0030076203402131796
step: 30, loss: 0.00579332048073411
step: 40, loss: 0.0011300276964902878
step: 50, loss: 0.02876497432589531
step: 60, loss: 0.010931489057838917
step: 70, loss: 0.0026711805257946253
step: 80, loss: 0.0020423950627446175
step: 90, loss: 0.002880264539271593
step: 100, loss: 0.03627670183777809
step: 110, loss: 0.022404033690690994
step: 120, loss: 0.08752122521400452
step: 130, loss: 0.0008807088015601039
step: 140, loss: 0.044439855962991714
step: 150, loss: 0.007734715007245541
step: 160, loss: 0.0011624748585745692
step: 170, loss: 0.002450151601806283
step: 180, loss: 0.0008830927545204759
step: 190, loss: 0.0007779984152875841
/usr/local/lib/python3.7/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=0.8465608465608466, f1=0.8631578947368422,
best_f1=0.851063829787234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018849496264010668
step: 10, loss: 0.15138423442840576
step: 20, loss: 0.0012699974467977881
step: 30, loss: 0.011173739098012447
step: 40, loss: 0.0008331039571203291
step: 50, loss: 0.0006926918285898864
step: 60, loss: 0.008137407712638378
step: 70, loss: 0.05366332456469536
step: 80, loss: 0.0008702035993337631
step: 90, loss: 0.0015167659148573875
step: 100, loss: 0.0003263706748839468
step: 110, loss: 0.0016104534734040499
step: 120, loss: 0.0015188846737146378
step: 130, loss: 0.0010880096815526485
step: 140, loss: 0.000796244537923485
step: 150, loss: 0.002263318281620741
step: 160, loss: 0.00257197511382401
step: 170, loss: 0.0004644645086955279
step: 180, loss: 0.0010757133131846786
step: 190, loss: 0.22879457473754883
/usr/local/lib/python3.7/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=0.8505154639175257, f1=0.8563968668407311,
best_f1=0.851063829787234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02142786979675293
step: 10, loss: 0.0012996660079807043
step: 20, loss: 0.0010900472989305854
step: 30, loss: 0.00298760156147182
step: 40, loss: 0.0009969896636903286
step: 50, loss: 0.000959906610660255
step: 60, loss: 0.0011641767341643572
step: 70, loss: 0.0011508556781336665
step: 80, loss: 0.0011726489756256342
step: 90, loss: 0.00044209312181919813
step: 100, loss: 0.0009916225681081414
step: 110, loss: 0.0010924385860562325
step: 120, loss: 0.0013037055032327771
step: 130, loss: 0.0015154265565797687
step: 140, loss: 0.0005684757488779724

```



```

step: 150, loss: 0.002242602873593569
step: 160, loss: 0.002033883472904563
step: 170, loss: 0.0006235353066585958
step: 180, loss: 0.0027266868855804205
step: 190, loss: 0.0020255702547729015
/usr/local/lib/python3.7/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=0.8548812664907651, f1=0.8525469168900803,
best_f1=0.851063829787234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029174680821597576
step: 10, loss: 0.0008042368572205305
step: 20, loss: 0.003189739538356662
step: 30, loss: 0.0008264358039014041
step: 40, loss: 0.0017161010764539242
step: 50, loss: 0.0007163392729125917
step: 60, loss: 0.0004977751523256302
step: 70, loss: 0.002705237828195095
step: 80, loss: 0.0006288858712650836
step: 90, loss: 0.0009244225220754743
step: 100, loss: 0.00103874655906111
step: 110, loss: 0.001611678977496922
step: 120, loss: 0.002733179833739996
step: 130, loss: 0.011770714074373245
step: 140, loss: 0.001188990892842412
step: 150, loss: 0.0005397865897975862
step: 160, loss: 0.0006440117140300572
step: 170, loss: 0.032955579459667206
step: 180, loss: 0.0009054081747308373
step: 190, loss: 0.0054138037376105785
/usr/local/lib/python3.7/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=0.851851851851852, f1=0.8525469168900803,
best_f1=0.851063829787234
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True

```

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:09, 211.82it/s]
load_f1 = 0.8548812664907651
real_f1 = 0.8496042216358839
733it [00:00, 3431.97it/s]/usr/local/lib/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, 169.91it/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 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: 0.49298232793807983
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.427386075258255
step: 20, loss: 0.3384810984134674
step: 30, loss: 0.4204258322715759
step: 40, loss: 0.6101604104042053
step: 50, loss: 0.3060830533504486
step: 60, loss: 0.6252105236053467
step: 70, loss: 0.3080914318561554
step: 80, loss: 0.23621247708797455
step: 90, loss: 0.18223486840724945
step: 100, loss: 0.17890338599681854
step: 110, loss: 0.35877367854118347
step: 120, loss: 0.3110013008117676
step: 130, loss: 0.30231228470802307
step: 140, loss: 0.3648180365562439
step: 150, loss: 0.38537028431892395
step: 160, loss: 0.41657930612564087
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.32944950461387634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.305885374546051
step: 10, loss: 0.4512103199958801
step: 20, loss: 0.31156137585639954
step: 30, loss: 0.33832183480262756
step: 40, loss: 0.09744970500469208
step: 50, loss: 0.4384151101112366
step: 60, loss: 0.20872588455677032
step: 70, loss: 0.5079758763313293
step: 80, loss: 0.24766911566257477
step: 90, loss: 0.23339645564556122
step: 100, loss: 0.5092775225639343
step: 110, loss: 0.25611627101898193
step: 120, loss: 0.26147499680519104

```

```

step: 130, loss: 0.5485310554504395
step: 140, loss: 0.5033047795295715
step: 150, loss: 0.4366288185119629
step: 160, loss: 0.4396890103816986
step: 170, loss: 0.3737252950668335
/usr/local/lib/python3.7/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=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5991235971450806
step: 10, loss: 0.3099208176136017
step: 20, loss: 0.24606813490390778
step: 30, loss: 0.24399639666080475
step: 40, loss: 0.37934553623199463
step: 50, loss: 0.5610774755477905
step: 60, loss: 0.31066131591796875
step: 70, loss: 0.23003429174423218
step: 80, loss: 0.3859483599662781
step: 90, loss: 0.52802574634552
step: 100, loss: 0.3093213140964508
step: 110, loss: 0.1700637936592102
step: 120, loss: 0.5940511226654053
step: 130, loss: 0.49736663699150085
step: 140, loss: 0.44444194436073303
step: 150, loss: 0.19481831789016724
step: 160, loss: 0.16743648052215576
step: 170, loss: 0.31128379702568054
/usr/local/lib/python3.7/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=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3836391866207123
step: 10, loss: 0.5298935770988464
step: 20, loss: 0.24600377678871155
step: 30, loss: 0.449623167514801
step: 40, loss: 0.2505738139152527
step: 50, loss: 0.3661210536956787
step: 60, loss: 0.6840435266494751
step: 70, loss: 0.32006850838661194
step: 80, loss: 0.4430324137210846

```

```

step: 90, loss: 0.30897918343544006
step: 100, loss: 0.3821682929992676
step: 110, loss: 0.452434241771698
step: 120, loss: 0.4949474632740021
step: 130, loss: 0.31483352184295654
step: 140, loss: 0.24751077592372894
step: 150, loss: 0.7078360915184021
step: 160, loss: 0.1253325641155243
step: 170, loss: 0.24223650991916656
/usr/local/lib/python3.7/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=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3844662606716156
step: 10, loss: 0.31177955865859985
step: 20, loss: 0.3253532350063324
step: 30, loss: 0.3148878812789917
step: 40, loss: 0.24114415049552917
step: 50, loss: 0.2426823377609253
step: 60, loss: 0.3105146288871765
step: 70, loss: 0.32453471422195435
step: 80, loss: 0.07843266427516937
step: 90, loss: 0.6312142014503479
step: 100, loss: 0.2618486285209656
step: 110, loss: 0.2594281733036041
step: 120, loss: 0.1381743848323822
step: 130, loss: 0.24318552017211914
step: 140, loss: 0.16529351472854614
step: 150, loss: 0.2895088791847229
step: 160, loss: 0.2515503466129303
step: 170, loss: 0.3694189488887787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2147165259348613, f1=0.22596754057428212,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3041476905345917

```

```

step: 10, loss: 0.09389384090900421
step: 20, loss: 0.36712828278541565
step: 30, loss: 0.24539348483085632
step: 40, loss: 0.4493888318538666
step: 50, loss: 0.2422264665365219
step: 60, loss: 0.38279619812965393
step: 70, loss: 0.38072240352630615
step: 80, loss: 0.20716263353824615
step: 90, loss: 0.2990924119949341
step: 100, loss: 0.17980323731899261
step: 110, loss: 0.44521307945251465
step: 120, loss: 0.448686808347702
step: 130, loss: 0.5577392578125
step: 140, loss: 0.22757495939731598
step: 150, loss: 0.16968567669391632
step: 160, loss: 0.3108125925064087
step: 170, loss: 0.2381090670824051
/usr/local/lib/python3.7/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=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23653574287891388
step: 10, loss: 0.4593958258628845
step: 20, loss: 0.46491295099258423
step: 30, loss: 0.514277994632721
step: 40, loss: 0.13357903063297272
step: 50, loss: 0.45204707980155945
step: 60, loss: 0.5437372922897339
step: 70, loss: 0.38233569264411926
step: 80, loss: 0.17115722596645355
step: 90, loss: 0.44978106021881104
step: 100, loss: 0.3126794695854187
step: 110, loss: 0.393180251121521
step: 120, loss: 0.3870450258255005
step: 130, loss: 0.19357141852378845
step: 140, loss: 0.12472053617238998
step: 150, loss: 0.316521555185318
step: 160, loss: 0.3814721405506134
step: 170, loss: 0.31341293454170227
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45186659693717957
step: 10, loss: 0.5235683917999268
step: 20, loss: 0.24643325805664062
step: 30, loss: 0.302647203207016
step: 40, loss: 0.23971597850322723
step: 50, loss: 0.3783785402774811
step: 60, loss: 0.37585291266441345
step: 70, loss: 0.19434291124343872
step: 80, loss: 0.3191175162792206
step: 90, loss: 0.57785964012146
step: 100, loss: 0.253431499004364
step: 110, loss: 0.50958251953125
step: 120, loss: 0.3173697590827942
step: 130, loss: 0.3176548182964325
step: 140, loss: 0.5094006657600403
step: 150, loss: 0.24825413525104523
step: 160, loss: 0.1840309351682663
step: 170, loss: 0.46289771795272827
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3815666735172272
step: 10, loss: 0.5801163911819458
step: 20, loss: 0.3777028024196625
step: 30, loss: 0.18228526413440704
step: 40, loss: 0.3840693235397339
step: 50, loss: 0.24968008697032928
step: 60, loss: 0.38453179597854614
step: 70, loss: 0.31325626373291016
step: 80, loss: 0.16304785013198853
step: 90, loss: 0.4550558030605316
step: 100, loss: 0.5057145953178406
step: 110, loss: 0.3122575581073761
step: 120, loss: 0.45334404706954956
step: 130, loss: 0.3167038559913635
step: 140, loss: 0.37770646810531616
step: 150, loss: 0.38444241881370544
step: 160, loss: 0.2435883730649948
step: 170, loss: 0.4395234286785126
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4446178376674652
step: 10, loss: 0.3068009912967682
step: 20, loss: 0.25563836097717285
step: 30, loss: 0.6497437953948975
step: 40, loss: 0.3655640482902527
step: 50, loss: 0.41994377970695496
step: 60, loss: 0.47957873344421387
step: 70, loss: 0.27502745389938354
step: 80, loss: 0.2561742067337036
step: 90, loss: 0.24045032262802124
step: 100, loss: 0.23573444783687592
step: 110, loss: 0.3130589723587036
step: 120, loss: 0.3111875653266907
step: 130, loss: 0.18418942391872406
step: 140, loss: 0.2393747866153717
step: 150, loss: 0.3141380548477173
step: 160, loss: 0.2518477439880371
step: 170, loss: 0.5153698325157166
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3748346269130707
step: 10, loss: 0.30665814876556396
step: 20, loss: 0.16894033551216125
step: 30, loss: 0.5244277119636536
step: 40, loss: 0.24169453978538513
step: 50, loss: 0.23118433356285095
step: 60, loss: 0.3767133951187134
step: 70, loss: 0.31982851028442383
step: 80, loss: 0.19032566249370575
step: 90, loss: 0.3872322738170624
step: 100, loss: 0.5277342200279236
step: 110, loss: 0.3788638114929199
step: 120, loss: 0.3685917556285858
step: 130, loss: 0.37719088792800903
step: 140, loss: 0.5062842965126038

```

```

step: 150, loss: 0.4340707063674927
step: 160, loss: 0.25385281443595886
step: 170, loss: 0.18278338015079498
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3883886933326721
step: 10, loss: 0.30761295557022095
step: 20, loss: 0.24377459287643433
step: 30, loss: 0.3102542757987976
step: 40, loss: 0.3101523220539093
step: 50, loss: 0.306840717792511
step: 60, loss: 0.2874014377593994
step: 70, loss: 0.19027113914489746
step: 80, loss: 0.244525745511055
step: 90, loss: 0.3141576051712036
step: 100, loss: 0.24994850158691406
step: 110, loss: 0.440424382686615
step: 120, loss: 0.3056800663471222
step: 130, loss: 0.3825685679912567
step: 140, loss: 0.30622661113739014
step: 150, loss: 0.3103830814361572
step: 160, loss: 0.6555073857307434
step: 170, loss: 0.3774455189704895
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.378624826669693
step: 10, loss: 0.25651589035987854
step: 20, loss: 0.18224136531352997
step: 30, loss: 0.4456676244735718
step: 40, loss: 0.38187891244888306
step: 50, loss: 0.24688315391540527
step: 60, loss: 0.2397581934928894
step: 70, loss: 0.38211631774902344
step: 80, loss: 0.24075093865394592
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.3097800016403198

```

```

step: 100, loss: 0.1777859330177307
step: 110, loss: 0.45015811920166016
step: 120, loss: 0.23729482293128967
step: 130, loss: 0.24616892635822296
step: 140, loss: 0.45626819133758545
step: 150, loss: 0.37621060013771057
step: 160, loss: 0.31559526920318604
step: 170, loss: 0.30878815054893494
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37982019782066345
step: 10, loss: 0.23928602039813995
step: 20, loss: 0.3806917369365692
step: 30, loss: 0.4457464814186096
step: 40, loss: 0.24452222883701324
step: 50, loss: 0.30320584774017334
step: 60, loss: 0.3102090656757355
step: 70, loss: 0.31158047914505005
step: 80, loss: 0.1831684112548828
step: 90, loss: 0.5051988363265991
step: 100, loss: 0.37970733642578125
step: 110, loss: 0.24595490097999573
step: 120, loss: 0.24854891002178192
step: 130, loss: 0.44565436244010925
step: 140, loss: 0.3813803791999817
step: 150, loss: 0.24168618023395538
step: 160, loss: 0.24413488805294037
step: 170, loss: 0.31492510437965393
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11145684123039246
step: 10, loss: 0.44769173860549927
step: 20, loss: 0.3131098747253418
step: 30, loss: 0.5161538124084473
step: 40, loss: 0.17950807511806488
step: 50, loss: 0.2485385239124298

```

```

step: 60, loss: 0.37830084562301636
step: 70, loss: 0.18021774291992188
step: 80, loss: 0.3124399185180664
step: 90, loss: 0.31744709610939026
step: 100, loss: 0.5138290524482727
step: 110, loss: 0.5721276998519897
step: 120, loss: 0.31661710143089294
step: 130, loss: 0.4427967369556427
step: 140, loss: 0.37121883034706116
step: 150, loss: 0.17157593369483948
step: 160, loss: 0.24397563934326172
step: 170, loss: 0.30790698528289795
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.22596754057428212
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:05, 319.46it/s]

```

```
load_f1 = 0.22418478260869565
real_f1 = 0.21449970041941283
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, 199.03it/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 was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5871965289115906
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.5000245571136475
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.5008035898208618
step: 30, loss: 0.29996198415756226
step: 40, loss: 0.387540340423584
step: 50, loss: 0.4297231435775757
step: 60, loss: 0.36345404386520386
step: 70, loss: 0.11580298840999603
step: 80, loss: 0.02705357037484646
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.32603684067726135
step: 100, loss: 0.11671431362628937
step: 110, loss: 0.03417884558439255
step: 120, loss: 0.10094451159238815
step: 130, loss: 0.25994688272476196
step: 140, loss: 0.12482935935258865
step: 150, loss: 0.20919759571552277
step: 160, loss: 0.010118572972714901
step: 170, loss: 0.17648768424987793
step: 180, loss: 0.09261544048786163
step: 190, loss: 0.01335029024630785
step: 200, loss: 0.023663364350795746
step: 210, loss: 0.030211305245757103
step: 220, loss: 0.12476040422916412
step: 230, loss: 0.006621276959776878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9582417582417582, f1=0.9631284916201117,
best_f1=0.9631284916201117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007188566029071808
step: 10, loss: 0.1496201604604721
step: 20, loss: 0.017606113106012344
step: 30, loss: 0.07130765169858932
step: 40, loss: 0.05555759742856026
step: 50, loss: 0.04726550355553627
step: 60, loss: 0.01172832865267992
step: 70, loss: 0.023125311359763145
step: 80, loss: 0.004007386974990368
step: 90, loss: 0.0068111727014184
step: 100, loss: 0.01591062732040882
step: 110, loss: 0.049661532044410706
step: 120, loss: 0.007036034017801285
step: 130, loss: 0.012905864976346493
step: 140, loss: 0.03996225818991661
step: 150, loss: 0.22423778474330902
step: 160, loss: 0.05157151818275452
step: 170, loss: 0.002652907744050026
step: 180, loss: 0.018875522539019585
step: 190, loss: 0.08300244808197021
step: 200, loss: 0.020721781998872757
step: 210, loss: 0.08244933933019638
step: 220, loss: 0.012049398384988308
step: 230, loss: 0.003109375014901161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9690265486725664, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0378086119890213

```

```

step: 10, loss: 0.006485857535153627
step: 20, loss: 0.002451385138556361
step: 30, loss: 0.0022694782819598913
step: 40, loss: 0.11364082247018814
step: 50, loss: 0.026858188211917877
step: 60, loss: 0.011708805337548256
step: 70, loss: 0.013113465160131454
step: 80, loss: 0.12679855525493622
step: 90, loss: 0.01632877066731453
step: 100, loss: 0.006474374793469906
step: 110, loss: 0.04782070964574814
step: 120, loss: 0.008576474152505398
step: 130, loss: 0.0071395509876310825
step: 140, loss: 0.002147712977603078
step: 150, loss: 0.07180926948785782
step: 160, loss: 0.006835007108747959
step: 170, loss: 0.003353276289999485
step: 180, loss: 0.023705149069428444
step: 190, loss: 0.02104816772043705
step: 200, loss: 0.007212997879832983
step: 210, loss: 0.0023660564329475164
step: 220, loss: 0.09084608405828476
step: 230, loss: 0.01912122592329979
/usr/local/lib/python3.7/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=0.9678135405105438, f1=0.9579646017699115,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009065763093531132
step: 10, loss: 0.0026390228886157274
step: 20, loss: 0.0059850928373634815
step: 30, loss: 0.0030743160750716925
step: 40, loss: 0.046450626105070114
step: 50, loss: 0.012808970175683498
step: 60, loss: 0.020895112305879593
step: 70, loss: 0.014992897398769855
step: 80, loss: 0.11635620892047882
step: 90, loss: 0.02874012105166912
step: 100, loss: 0.011813932098448277
step: 110, loss: 0.005605127662420273
step: 120, loss: 0.018707869574427605
step: 130, loss: 0.012129292823374271
step: 140, loss: 0.0037806371692568064
step: 150, loss: 0.005184893496334553
step: 160, loss: 0.008692513220012188

```



```

step: 170, loss: 0.0029919587541371584
step: 180, loss: 0.08100325614213943
step: 190, loss: 0.010875653475522995
step: 200, loss: 0.08889537304639816
step: 210, loss: 0.01575182005763054
step: 220, loss: 0.002575457561761141
step: 230, loss: 0.007945910096168518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9755011135857461, f1=0.96875, best_f1=0.96875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019521925132721663
step: 10, loss: 0.002927385037764907
step: 20, loss: 0.15618830919265747
step: 30, loss: 0.012918546795845032
step: 40, loss: 0.013810576871037483
step: 50, loss: 0.001049507176503539
step: 60, loss: 0.014883341267704964
step: 70, loss: 0.011989622376859188
step: 80, loss: 0.012181010097265244
step: 90, loss: 0.21359868347644806
step: 100, loss: 0.0015728144207969308
step: 110, loss: 0.005946779157966375
step: 120, loss: 0.003253244562074542
step: 130, loss: 0.0006281978567130864
step: 140, loss: 0.001144134090282023
step: 150, loss: 0.01389290951192379
step: 160, loss: 0.0005955703090876341
step: 170, loss: 0.011142105795443058
step: 180, loss: 0.003320799209177494
step: 190, loss: 0.14061041176319122
step: 200, loss: 0.0707918256521225
step: 210, loss: 0.003921295516192913
step: 220, loss: 0.006325291004031897
step: 230, loss: 0.07691695541143417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9698324022346367,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004305437207221985
step: 10, loss: 0.0024950741790235043
step: 20, loss: 0.004066964145749807
step: 30, loss: 0.005713955033570528
step: 40, loss: 0.0021464393939822912
step: 50, loss: 0.0057816701009869576
step: 60, loss: 0.0026049709413200617
step: 70, loss: 0.18363448977470398
step: 80, loss: 0.0016843607882037759
step: 90, loss: 0.004254082217812538
step: 100, loss: 0.004036797676235437
step: 110, loss: 0.02529010735452175
step: 120, loss: 0.0020705980714410543
step: 130, loss: 0.007722147740423679
step: 140, loss: 0.0028605060651898384
step: 150, loss: 0.0004053676675539464
step: 160, loss: 0.015040713362395763
step: 170, loss: 0.0011300306068733335
step: 180, loss: 0.0012925693299621344
step: 190, loss: 0.006232079584151506
step: 200, loss: 0.004310097079724073
step: 210, loss: 0.0020917255897074938
step: 220, loss: 0.001430393778719008
step: 230, loss: 0.001958745764568448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.980963045912654, 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: 0.004410304129123688
step: 10, loss: 0.006291416473686695
step: 20, loss: 0.006318996660411358
step: 30, loss: 0.005141150671988726
step: 40, loss: 0.0010524122044444084
step: 50, loss: 0.0017498821252956986
step: 60, loss: 0.005408206023275852

```

```

step: 70, loss: 0.0007162751280702651
step: 80, loss: 0.0016913912259042263
step: 90, loss: 0.0022978992201387882
step: 100, loss: 0.0014887358993291855
step: 110, loss: 0.0007053057197481394
step: 120, loss: 0.018373508006334305
step: 130, loss: 0.0006512411637231708
step: 140, loss: 0.00034870431409217417
step: 150, loss: 0.18153265118598938
step: 160, loss: 0.0007325756596401334
step: 170, loss: 0.002502795308828354
step: 180, loss: 0.0004935358883813024
step: 190, loss: 0.0005359251517802477
step: 200, loss: 0.0020231110975146294
step: 210, loss: 0.001072992687113583
step: 220, loss: 0.0017319949110969901
step: 230, loss: 0.002878764644265175
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, f1=0.9841269841269841,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003323900979012251
step: 10, loss: 0.008729864843189716
step: 20, loss: 0.003754468634724617
step: 30, loss: 0.0035086756106466055
step: 40, loss: 0.0013720211572945118
step: 50, loss: 0.000988081330433488
step: 60, loss: 0.0007063959492370486
step: 70, loss: 0.00039597158320248127
step: 80, loss: 0.012701272964477539
step: 90, loss: 0.00044110670569352806
step: 100, loss: 0.0008504111901856959
step: 110, loss: 0.0006537455483339727
step: 120, loss: 0.0017014896729961038
step: 130, loss: 0.0031831618398427963
step: 140, loss: 0.00040814862586557865
step: 150, loss: 0.15723372995853424
step: 160, loss: 0.00024015235248953104
step: 170, loss: 0.002542440313845873
step: 180, loss: 0.004066627938300371

```

```

step: 190, loss: 0.0003949832753278315
step: 200, loss: 0.03515033796429634
step: 210, loss: 0.0010833413107320666
step: 220, loss: 0.000980574986897409
step: 230, loss: 0.0006193366134539247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854096520763187, 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: 0.0005564430612139404
step: 10, loss: 0.0002900762774515897
step: 20, loss: 0.00024934179964475334
step: 30, loss: 0.0003831575158983469
step: 40, loss: 0.00042485943413339555
step: 50, loss: 0.0006624468369409442
step: 60, loss: 0.0008584560710005462
step: 70, loss: 0.08424080163240433
step: 80, loss: 0.0002620238810777664
step: 90, loss: 0.028458334505558014
step: 100, loss: 0.009107396937906742
step: 110, loss: 0.000348731002304703
step: 120, loss: 0.05780446156859398
step: 130, loss: 0.1710529625415802
step: 140, loss: 0.002488724421709776
step: 150, loss: 0.00121983140707016
step: 160, loss: 0.005760674364864826
step: 170, loss: 0.00038210986531339586
step: 180, loss: 0.005161644890904427
step: 190, loss: 0.00026302991318516433
step: 200, loss: 0.0008597932755947113
step: 210, loss: 0.009381430223584175
step: 220, loss: 0.0030955690890550613
step: 230, loss: 0.00023424989194609225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887133182844244, f1=0.9841269841269841,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011546696769073606
step: 10, loss: 0.00040749451727606356
step: 20, loss: 0.00037960524787195027
step: 30, loss: 0.0001938642089953646
step: 40, loss: 0.006335355341434479
step: 50, loss: 0.00017737096641212702
step: 60, loss: 0.0003245798870921135
step: 70, loss: 0.006381392013281584
step: 80, loss: 0.0002305733651155606
step: 90, loss: 0.00018563451885711402
step: 100, loss: 0.0003368591715116054
step: 110, loss: 0.00039062160067260265
step: 120, loss: 0.00029989040922373533
step: 130, loss: 0.0006881315493956208
step: 140, loss: 0.00017028175352606922
step: 150, loss: 0.0006215384928509593
step: 160, loss: 7.023617217782885e-05
step: 170, loss: 0.00047096304479055107
step: 180, loss: 0.0008219671435654163
step: 190, loss: 0.0006936213467270136
step: 200, loss: 0.00044310803059488535
step: 210, loss: 0.0010735468240454793
step: 220, loss: 0.008749528788030148
step: 230, loss: 0.00040804242598824203
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9795454545454545,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015277770580723882
step: 10, loss: 0.0006133975693956017
step: 20, loss: 0.0011029638117179275
step: 30, loss: 0.005052154418081045
step: 40, loss: 0.000562510802410543
step: 50, loss: 0.006771252956241369
step: 60, loss: 0.014105512760579586

```

```

step: 70, loss: 0.002116844290867448
step: 80, loss: 0.000652506947517395
step: 90, loss: 0.0036864166613668203
step: 100, loss: 0.0003413253871258348
step: 110, loss: 0.006285320967435837
step: 120, loss: 0.0002479629765730351
step: 130, loss: 0.0002591335214674473
step: 140, loss: 0.007531685288995504
step: 150, loss: 0.0002741591597441584
step: 160, loss: 0.09150578081607819
step: 170, loss: 0.0039461106061935425
step: 180, loss: 0.0016569190192967653
step: 190, loss: 0.0010733706876635551
step: 200, loss: 0.004168349783867598
step: 210, loss: 0.0003955601423513144
step: 220, loss: 0.0003690526937134564
step: 230, loss: 0.0001857689203461632
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9796839729119639,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00044950511073693633
step: 10, loss: 0.0001393697311868891
step: 20, loss: 0.009692267514765263
step: 30, loss: 0.03271285817027092
step: 40, loss: 0.0008719887118786573
step: 50, loss: 0.0005641502211801708
step: 60, loss: 0.0005658676382154226
step: 70, loss: 0.00018997739243786782
step: 80, loss: 7.915677269920707e-05
step: 90, loss: 0.00040582913788966835
step: 100, loss: 0.00014285187353380024
step: 110, loss: 0.0001179134487756528
step: 120, loss: 0.00039084936724975705
step: 130, loss: 0.00034949934342876077
step: 140, loss: 0.0002362837258260697
step: 150, loss: 0.0008696695440448821
step: 160, loss: 0.0003664073010440916
step: 170, loss: 0.0004752856330014765
step: 180, loss: 0.00037873288965784013
step: 190, loss: 0.0004193955101072788
step: 200, loss: 0.0002182823373004794
step: 210, loss: 0.008495702408254147
step: 220, loss: 0.006122162565588951

```

```

step: 230, loss: 0.0028773630037903786
/usr/local/lib/python3.7/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=0.9887640449438202, f1=0.9796839729119639,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001143603352829814
step: 10, loss: 0.005046777427196503
step: 20, loss: 0.00020229317306075245
step: 30, loss: 0.00014195824041962624
step: 40, loss: 0.0003194189630448818
step: 50, loss: 0.0004965533735230565
step: 60, loss: 0.00019500087364576757
step: 70, loss: 0.0012655970640480518
step: 80, loss: 0.00011279738828307018
step: 90, loss: 0.001233246992342174
step: 100, loss: 0.0001709365751594305
step: 110, loss: 0.00022021769837010652
step: 120, loss: 0.0002204058546340093
step: 130, loss: 0.00028074768488295376
step: 140, loss: 0.00041713929385878146
step: 150, loss: 0.0005722020869143307
step: 160, loss: 0.0008257631561718881
step: 170, loss: 0.00018069767975248396
step: 180, loss: 0.025233056396245956
step: 190, loss: 0.00026004441315308213
step: 200, loss: 5.439031156129204e-05
step: 210, loss: 0.00018704778631217778
step: 220, loss: 0.0001416869490640238
step: 230, loss: 7.776049460517243e-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=0.9898534385569334, f1=0.9840546697038726,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010057345207314938
step: 10, loss: 0.00013851455878466368
step: 20, loss: 0.00010285422467859462
step: 30, loss: 0.0001628737081773579
step: 40, loss: 8.746064122533426e-05
step: 50, loss: 8.646195055916905e-05
step: 60, loss: 0.00015184022777248174

```

```

step: 70, loss: 8.935183723224327e-05
step: 80, loss: 0.0001566940627526492
step: 90, loss: 0.00014968914911150932
step: 100, loss: 8.872356556821615e-05
step: 110, loss: 0.00012794342183042318
step: 120, loss: 3.44999716617167e-05
step: 130, loss: 0.00016470668197143823
step: 140, loss: 0.00010433256829855964
step: 150, loss: 3.663827737909742e-05
step: 160, loss: 8.91359886736609e-05
step: 170, loss: 0.00024395022774115205
step: 180, loss: 0.00010049852426163852
step: 190, loss: 0.00012361211702227592
step: 200, loss: 0.00019928955589421093
step: 210, loss: 5.865876664756797e-05
step: 220, loss: 0.00024414705694653094
step: 230, loss: 0.0015415430534631014
/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9829738933030647,
best_f1=0.9795454545454545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011732527054846287
step: 10, loss: 0.00021205347729846835
step: 20, loss: 0.0004823749477509409
step: 30, loss: 0.000182965217391029
step: 40, loss: 4.859473847318441e-05
step: 50, loss: 6.231141014723107e-05
step: 60, loss: 0.020830994471907616
step: 70, loss: 8.641216845717281e-05
step: 80, loss: 8.397086639888585e-05
step: 90, loss: 0.00010769529035314918
step: 100, loss: 7.815729622961953e-05
step: 110, loss: 9.560170292388648e-05
step: 120, loss: 0.03185899555683136
step: 130, loss: 0.0001714274985715747
step: 140, loss: 0.02871854230761528
step: 150, loss: 0.0001943591341841966
step: 160, loss: 0.000660564168356359
step: 170, loss: 4.458483817870729e-05
step: 180, loss: 0.0001994256308535114
step: 190, loss: 0.0001459712366340682
step: 200, loss: 0.003962567541748285
step: 210, loss: 0.007575906813144684
step: 220, loss: 7.73467036196962e-05

```



```

step: 230, loss: 0.0001395174185745418
/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9829738933030647,
best_f1=0.9795454545454545
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 210.98it/s]
load_f1 = 0.9887133182844244
real_f1 = 0.9898534385569334
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, 193.96it/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: 0.617305338382721
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: 0.42008450627326965
step: 20, loss: 0.2577465772628784
step: 30, loss: 0.3543079197406769
step: 40, loss: 0.3730810582637787
step: 50, loss: 0.5418809652328491
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.20696187019348145
step: 70, loss: 0.19423460960388184
step: 80, loss: 0.18502813577651978
step: 90, loss: 0.18483638763427734
step: 100, loss: 0.297818660736084
step: 110, loss: 0.07054140418767929
step: 120, loss: 0.17044156789779663
step: 130, loss: 0.17726284265518188
step: 140, loss: 0.26821428537368774
step: 150, loss: 0.14254213869571686
step: 160, loss: 0.2340957075357437
step: 170, loss: 0.14892710745334625
step: 180, loss: 0.0946924239397049
step: 190, loss: 0.04963673651218414
step: 200, loss: 0.03294730558991432
step: 210, loss: 0.03931895270943642
step: 220, loss: 0.10126622766256332
step: 230, loss: 0.15789847075939178
step: 240, loss: 0.05125819146633148
step: 250, loss: 0.05727346986532211
step: 260, loss: 0.2504291236400604
step: 270, loss: 0.42372530698776245
step: 280, loss: 0.05882978439331055
step: 290, loss: 0.06903392821550369
step: 300, loss: 0.17785993218421936
step: 310, loss: 0.16980887949466705
step: 320, loss: 0.09022986143827438
step: 330, loss: 0.05139005184173584
step: 340, loss: 0.4485205113887787
step: 350, loss: 0.09307119995355606

```

```

step: 360, loss: 0.0861373320221901
step: 370, loss: 0.030184917151927948
step: 380, loss: 0.15786443650722504
step: 390, loss: 0.013369842432439327
step: 400, loss: 0.019190222024917603
step: 410, loss: 0.2810988128185272
step: 420, loss: 0.010698881931602955
step: 430, loss: 0.005085306242108345
step: 440, loss: 0.02768009714782238
step: 450, loss: 0.021685592830181122
step: 460, loss: 0.012165584601461887
step: 470, loss: 0.025463668629527092
step: 480, loss: 0.17987947165966034
step: 490, loss: 0.11236995458602905
step: 500, loss: 0.019150501117110252
step: 510, loss: 0.015806149691343307
step: 520, loss: 0.06394055485725403
step: 530, loss: 0.004985817242413759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.937995337995338, f1=0.9380530973451328,
best_f1=0.9380530973451328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04321742057800293
step: 10, loss: 0.06271646916866302
step: 20, loss: 0.05520421639084816
step: 30, loss: 0.0526004433631897
step: 40, loss: 0.03485271334648132
step: 50, loss: 0.03197628632187843
step: 60, loss: 0.040362633764743805
step: 70, loss: 0.005234548356384039
step: 80, loss: 0.018895013257861137
step: 90, loss: 0.012057206593453884
step: 100, loss: 0.16761940717697144
step: 110, loss: 0.012381263077259064
step: 120, loss: 0.056702371686697006
step: 130, loss: 0.030302023515105247
step: 140, loss: 0.07271908968687057
step: 150, loss: 0.010771572589874268
step: 160, loss: 0.045588862150907516
step: 170, loss: 0.015865035355091095

```

```

step: 180, loss: 0.053533222526311874
step: 190, loss: 0.00551432091742754
step: 200, loss: 0.03558233007788658
step: 210, loss: 0.061455074697732925
step: 220, loss: 0.0013491198187693954
step: 230, loss: 0.030358852818608284
step: 240, loss: 0.17038790881633759
step: 250, loss: 0.005212245509028435
step: 260, loss: 0.07603900134563446
step: 270, loss: 0.013902590610086918
step: 280, loss: 0.018102597445249557
step: 290, loss: 0.033053476363420486
step: 300, loss: 0.052700210362672806
step: 310, loss: 0.02142520807683468
step: 320, loss: 0.01334219891577959
step: 330, loss: 0.06942152231931686
step: 340, loss: 0.09883175045251846
step: 350, loss: 0.004719214513897896
step: 360, loss: 0.13821806013584137
step: 370, loss: 0.021542100235819817
step: 380, loss: 0.1594780832529068
step: 390, loss: 0.011240009218454361
step: 400, loss: 0.09555181860923767
step: 410, loss: 0.023882191628217697
step: 420, loss: 0.042962994426488876
step: 430, loss: 0.34174618124961853
step: 440, loss: 0.017439650371670723
step: 450, loss: 0.03097829595208168
step: 460, loss: 0.10680239647626877
step: 470, loss: 0.011489762924611568
step: 480, loss: 0.02069885842502117
step: 490, loss: 0.015233644284307957
step: 500, loss: 0.005120719317346811
step: 510, loss: 0.021029267460107803
step: 520, loss: 0.294346421957016
step: 530, loss: 0.11740688234567642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9522497704315885, f1=0.9464450600184672,
best_f1=0.9464450600184672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.13593286275863647
step: 10, loss: 0.008156116120517254
step: 20, loss: 0.01042169239372015
step: 30, loss: 0.023084254935383797
step: 40, loss: 0.04573167487978935
step: 50, loss: 0.003633409272879362
step: 60, loss: 0.023199208080768585
step: 70, loss: 0.004561202600598335
step: 80, loss: 0.060067299753427505
step: 90, loss: 0.009400340728461742
step: 100, loss: 0.04495793208479881
step: 110, loss: 0.050719212740659714
step: 120, loss: 0.08336349576711655
step: 130, loss: 0.008783572353422642
step: 140, loss: 0.016383208334445953
step: 150, loss: 0.012432409450411797
step: 160, loss: 0.050834354013204575
step: 170, loss: 0.0392475351691246
step: 180, loss: 0.005115864332765341
step: 190, loss: 0.0022693921346217394
step: 200, loss: 0.008123529143631458
step: 210, loss: 0.003532853676006198
step: 220, loss: 0.046402063220739365
step: 230, loss: 0.011064683087170124
step: 240, loss: 0.05884207785129547
step: 250, loss: 0.055149197578430176
step: 260, loss: 0.12537957727909088
step: 270, loss: 0.0037784345913678408
step: 280, loss: 0.005954150576144457
step: 290, loss: 0.014207691885530949
step: 300, loss: 0.11940436065196991
step: 310, loss: 0.09202983230352402
step: 320, loss: 0.006261357571929693
step: 330, loss: 0.014443375170230865
step: 340, loss: 0.0039989701472222805
step: 350, loss: 0.14501811563968658
step: 360, loss: 0.005547653883695602
step: 370, loss: 0.040691930800676346
step: 380, loss: 0.006050597410649061
step: 390, loss: 0.006593446712940931
step: 400, loss: 0.04788361117243767
step: 410, loss: 0.15046025812625885
step: 420, loss: 0.005871338304132223
step: 430, loss: 0.007277047261595726
step: 440, loss: 0.1849880814552307
step: 450, loss: 0.06894563138484955
step: 460, loss: 0.09393006563186646
step: 470, loss: 0.03560304269194603

```

step: 480, loss: 0.14749644696712494
step: 490, loss: 0.0120299207046628
step: 500, loss: 0.013479567132890224
step: 510, loss: 0.007955639623105526
step: 520, loss: 0.0031004268676042557
step: 530, loss: 0.011378894560039043
/usr/local/lib/python3.7/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=0.9495136637332099, f1=0.9386046511627907,
best_f1=0.9464450600184672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02092263661324978
step: 10, loss: 0.0034722695127129555
step: 20, loss: 0.008423364721238613
step: 30, loss: 0.13569535315036774
step: 40, loss: 0.005634138360619545
step: 50, loss: 0.011143376119434834
step: 60, loss: 0.0030083719175308943
step: 70, loss: 0.017551083117723465
step: 80, loss: 0.14496512711048126
step: 90, loss: 0.04974616691470146
step: 100, loss: 0.0045690289698541164
step: 110, loss: 0.07301796972751617
step: 120, loss: 0.00147174671292305
step: 130, loss: 0.034224413335323334
step: 140, loss: 0.10296383500099182
step: 150, loss: 0.014282557182013988
step: 160, loss: 0.01741110347211361
step: 170, loss: 0.012548794038593769
step: 180, loss: 0.15327467024326324
step: 190, loss: 0.030660521239042282
step: 200, loss: 0.14564645290374756
step: 210, loss: 0.0007655073422938585
step: 220, loss: 0.00227446760982275
step: 230, loss: 0.0025439998134970665
step: 240, loss: 0.017359422519803047
step: 250, loss: 0.12035421282052994
step: 260, loss: 0.09837653487920761
step: 270, loss: 0.1323380321264267
step: 280, loss: 0.007321150507777929
step: 290, loss: 0.03522041440010071
step: 300, loss: 0.004842978902161121
step: 310, loss: 0.003965811803936958
step: 320, loss: 0.07872701436281204
step: 330, loss: 0.02173459716141224

```

```

step: 340, loss: 0.008723549544811249
step: 350, loss: 0.013698612339794636
step: 360, loss: 0.10538779944181442
step: 370, loss: 0.007158169522881508
step: 380, loss: 0.0071687353774905205
step: 390, loss: 0.001392711652442813
step: 400, loss: 0.00845019705593586
step: 410, loss: 0.0023587446194142103
step: 420, loss: 0.022761046886444092
step: 430, loss: 0.014781158417463303
step: 440, loss: 0.001483715372160077
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 450, loss: 0.03294163942337036
step: 460, loss: 0.0422995388507843
step: 470, loss: 0.003136781742796302
step: 480, loss: 0.05302049592137337
step: 490, loss: 0.0061478326097130775
step: 500, loss: 0.025093184784054756
step: 510, loss: 0.03960813954472542
step: 520, loss: 0.015775438398122787
step: 530, loss: 0.1498325616121292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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=0.9609665427509295, f1=0.9506517690875234,
best_f1=0.9506517690875234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013197464868426323
step: 10, loss: 0.011448467150330544
step: 20, loss: 0.01583755761384964
step: 30, loss: 0.02288072369992733
step: 40, loss: 0.003204931737855077
step: 50, loss: 0.08868569135665894
step: 60, loss: 0.013872516341507435
step: 70, loss: 0.00536866020411253
step: 80, loss: 0.00024151102115865797
step: 90, loss: 0.02208477072417736
step: 100, loss: 0.14764238893985748
step: 110, loss: 0.002740978030487895
step: 120, loss: 0.22174972295761108
step: 130, loss: 0.01035134308040142
step: 140, loss: 0.032480306923389435

```



```

step: 150, loss: 0.006246306002140045
step: 160, loss: 0.01641000621020794
step: 170, loss: 0.07516045868396759
step: 180, loss: 0.009099392220377922
step: 190, loss: 0.006793331354856491
step: 200, loss: 0.009745280258357525
step: 210, loss: 0.0015890150098130107
step: 220, loss: 0.002185167046263814
step: 230, loss: 0.0045736185275018215
step: 240, loss: 0.0014926525764167309
step: 250, loss: 0.21833111345767975
step: 260, loss: 0.0012496408307924867
step: 270, loss: 0.1467735916376114
step: 280, loss: 0.030946340411901474
step: 290, loss: 0.007787235081195831
step: 300, loss: 0.04880271479487419
step: 310, loss: 0.05868791043758392
step: 320, loss: 0.11766023188829422
step: 330, loss: 0.000704047386534512
step: 340, loss: 0.00568791851401329
step: 350, loss: 0.0008806160185486078
step: 360, loss: 0.0011874140473082662
step: 370, loss: 0.002421117853373289
step: 380, loss: 0.0012971542309969664
step: 390, loss: 0.06199254095554352
step: 400, loss: 0.004425797611474991
step: 410, loss: 0.06110404059290886
step: 420, loss: 0.21791768074035645
step: 430, loss: 0.008114061318337917
step: 440, loss: 0.000991045730188489
step: 450, loss: 0.004060983657836914
step: 460, loss: 0.011040298268198967
step: 470, loss: 0.04309249296784401
step: 480, loss: 0.0056206341832876205
step: 490, loss: 0.005060310009866953
step: 500, loss: 0.00792383961379528
step: 510, loss: 0.0008343790541402996
step: 520, loss: 0.2605152428150177
step: 530, loss: 0.03240272030234337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9614325068870523, f1=0.9529085872576176,
best_f1=0.9529085872576176

```

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

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

```
step: 0, loss: 0.014473228715360165  
step: 10, loss: 0.001689750817604363  
step: 20, loss: 0.000671888526994735  
step: 30, loss: 0.00276482617482543  
step: 40, loss: 0.004186505917459726  
step: 50, loss: 0.0008320813067257404  
step: 60, loss: 0.0017479460220783949  
step: 70, loss: 0.0028210943564772606  
step: 80, loss: 0.0012960645835846663  
step: 90, loss: 0.0006830082274973392  
step: 100, loss: 0.07049108296632767  
step: 110, loss: 0.0006556350272148848  
step: 120, loss: 0.0023880130611360073  
step: 130, loss: 0.0015119453892111778  
step: 140, loss: 0.0008225392084568739  
step: 150, loss: 0.0004880010092165321  
step: 160, loss: 0.06859974563121796  
step: 170, loss: 0.0014870238956063986  
step: 180, loss: 0.00048786040861159563  
step: 190, loss: 0.027382424101233482  
step: 200, loss: 0.002832782920449972  
step: 210, loss: 0.004687528591603041  
step: 220, loss: 0.004632893018424511  
step: 230, loss: 0.05355451628565788  
step: 240, loss: 0.00020563107682392  
step: 250, loss: 0.05932236090302467  
step: 260, loss: 0.00022685487056151032  
step: 270, loss: 0.0006171244313009083  
step: 280, loss: 0.00012546201469376683  
step: 290, loss: 0.0015023720916360617  
step: 300, loss: 0.004063871223479509  
step: 310, loss: 0.09324043244123459  
step: 320, loss: 0.002227658173069358  
step: 330, loss: 0.002647003624588251  
step: 340, loss: 0.002694563940167427  
step: 350, loss: 0.001343472977168858  
step: 360, loss: 0.07045935839414597  
step: 370, loss: 0.005900884047150612  
step: 380, loss: 0.0009654413443058729  
step: 390, loss: 0.006596451625227928  
step: 400, loss: 0.022801507264375687  
step: 410, loss: 0.0007382640615105629  
step: 420, loss: 0.0014862208627164364  
step: 430, loss: 0.00022549208370037377  
step: 440, loss: 0.0003939890011679381
```

```

step: 450, loss: 0.13933725655078888
step: 460, loss: 0.005805083084851503
step: 470, loss: 0.0007665801094844937
step: 480, loss: 0.0037752026692032814
step: 490, loss: 0.007340676616877317
step: 500, loss: 0.0017179729184135795
step: 510, loss: 0.2512204647064209
step: 520, loss: 0.0021192855201661587
step: 530, loss: 0.01689610257744789
/usr/local/lib/python3.7/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=0.9494949494949496, f1=0.9450853714813106,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0052237920463085175
step: 10, loss: 0.0007159491069614887
step: 20, loss: 0.0006678436184301972
step: 30, loss: 0.02332254871726036
step: 40, loss: 0.0016553145833313465
step: 50, loss: 0.011144855059683323
step: 60, loss: 0.0017533347709104419
step: 70, loss: 0.0005883611156605184
step: 80, loss: 0.0005573654198087752
step: 90, loss: 0.00019894390425179154
step: 100, loss: 0.17949560284614563
step: 110, loss: 0.0003191329597029835
step: 120, loss: 0.0027759037911891937
step: 130, loss: 0.00013132662570569664
step: 140, loss: 0.0009695534827187657
step: 150, loss: 0.0009255257318727672
step: 160, loss: 0.0001507766719441861
step: 170, loss: 0.003612207481637597
step: 180, loss: 0.002940525533631444
step: 190, loss: 0.003317685564979911
step: 200, loss: 0.002212563995271921
step: 210, loss: 0.000905283959582448
step: 220, loss: 0.00017923471750691533
step: 230, loss: 0.00043160567292943597
step: 240, loss: 0.002221053931862116
step: 250, loss: 0.0027405715081840754
step: 260, loss: 0.0029729013331234455
step: 270, loss: 0.002987445332109928
step: 280, loss: 0.010181878693401814
step: 290, loss: 0.016348110511898994
step: 300, loss: 0.000643097038846463

```

```

step: 310, loss: 0.0009217565529979765
step: 320, loss: 0.0006827613105997443
step: 330, loss: 0.0003678989305626601
step: 340, loss: 0.004493289161473513
step: 350, loss: 0.00014007551362738013
step: 360, loss: 0.003143847454339266
step: 370, loss: 0.006513169035315514
step: 380, loss: 0.00220590946264565
step: 390, loss: 0.01666136644780636
step: 400, loss: 0.00660294434055686
step: 410, loss: 0.002527235308662057
step: 420, loss: 0.030503256246447563
step: 430, loss: 0.0001154004639829509
step: 440, loss: 0.0008918349049054086
step: 450, loss: 0.0010985188418999314
step: 460, loss: 0.005886183585971594
step: 470, loss: 0.2721446454524994
step: 480, loss: 0.012210375629365444
step: 490, loss: 0.00474375719204545
step: 500, loss: 0.0023159158881753683
step: 510, loss: 0.0005935208755545318
step: 520, loss: 0.0012117082951590419
step: 530, loss: 0.007696259301155806
/usr/local/lib/python3.7/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=0.9548627268496975, f1=0.9501630181648812,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008317728061228991
step: 10, loss: 0.0016602183459326625
step: 20, loss: 0.0007918974151834846
step: 30, loss: 0.008731961250305176
step: 40, loss: 0.00018261015065945685
step: 50, loss: 0.002789704827591777
step: 60, loss: 0.0034465468488633633
step: 70, loss: 0.0008556615212000906
step: 80, loss: 0.0004519884823821485
step: 90, loss: 0.00047918304335325956
step: 100, loss: 0.00019865464128088206
step: 110, loss: 0.052052367478609085
step: 120, loss: 0.003960697911679745
step: 130, loss: 0.001323141623288393
step: 140, loss: 0.00016792611859273165
step: 150, loss: 0.0010257940739393234
step: 160, loss: 0.0005207304493524134

```

```

step: 170, loss: 0.09477640688419342
step: 180, loss: 0.0010371312964707613
step: 190, loss: 0.016350621357560158
step: 200, loss: 0.00444053765386343
step: 210, loss: 0.07575850188732147
step: 220, loss: 0.0009108471567742527
step: 230, loss: 0.12724335491657257
step: 240, loss: 0.0011752396821975708
step: 250, loss: 0.00011842857929877937
step: 260, loss: 0.00011261238978477195
step: 270, loss: 0.014419281855225563
step: 280, loss: 0.0006546757649630308
step: 290, loss: 0.004747520200908184
step: 300, loss: 7.315395487239584e-05
step: 310, loss: 0.012643365189433098
step: 320, loss: 0.0003134778526145965
step: 330, loss: 0.00016330776270478964
step: 340, loss: 0.0053735473193228245
step: 350, loss: 0.004605501424521208
step: 360, loss: 0.017241358757019043
step: 370, loss: 0.1019834578037262
step: 380, loss: 7.3743358370848e-05
step: 390, loss: 0.08276964724063873
step: 400, loss: 0.010756625793874264
step: 410, loss: 0.0003875444526784122
step: 420, loss: 0.0061208163388073444
step: 430, loss: 0.0017057110089808702
step: 440, loss: 0.0032887663692235947
step: 450, loss: 0.0052588945254683495
step: 460, loss: 0.0015606736997142434
step: 470, loss: 0.05642182379961014
step: 480, loss: 0.0003612485888879746
step: 490, loss: 0.006634044460952282
step: 500, loss: 0.001798318699002266
step: 510, loss: 0.018567318096756935
step: 520, loss: 7.569832814624533e-05
step: 530, loss: 0.025118835270404816
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 8: dev_f1=0.9560795191863153, f1=0.9463459759481962,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014650094090029597
step: 10, loss: 0.01834086887538433

```

step: 20, loss: 0.00011239697050768882
step: 30, loss: 0.11877986788749695
step: 40, loss: 0.002236902015283704
step: 50, loss: 0.0003999665204901248
step: 60, loss: 0.008420349098742008
step: 70, loss: 0.05992528051137924
step: 80, loss: 0.0005232241237536073
step: 90, loss: 0.059856753796339035
step: 100, loss: 0.0012319121742621064
step: 110, loss: 0.0024116733111441135
step: 120, loss: 0.038141921162605286
step: 130, loss: 0.0031733845826238394
step: 140, loss: 0.0035521488171070814
step: 150, loss: 0.005202615167945623
step: 160, loss: 0.0006524888449348509
step: 170, loss: 0.01077669020742178
step: 180, loss: 0.002933074254542589
step: 190, loss: 0.0005961129209026694
step: 200, loss: 0.007694531697779894
step: 210, loss: 0.0018111560493707657
step: 220, loss: 0.0020759019535034895
step: 230, loss: 0.00029134753276593983
step: 240, loss: 9.985187352867797e-05
step: 250, loss: 0.00017098644457291812
step: 260, loss: 0.0004203937132842839
step: 270, loss: 0.002594979014247656
step: 280, loss: 0.011184933595359325
step: 290, loss: 9.863284503808245e-05
step: 300, loss: 0.00011708238162100315
step: 310, loss: 0.0035291414242237806
step: 320, loss: 7.381673640338704e-05
step: 330, loss: 7.530390575993806e-05
step: 340, loss: 0.0007316778064705431
step: 350, loss: 0.0028888562228530645
step: 360, loss: 6.728825974278152e-05
step: 370, loss: 0.0004512644081842154
step: 380, loss: 0.013672851026058197
step: 390, loss: 0.0007038854528218508
step: 400, loss: 0.07736411690711975
step: 410, loss: 0.0012559759197756648
step: 420, loss: 9.479353320784867e-05
step: 430, loss: 0.029700519517064095
step: 440, loss: 0.00011053884372813627
step: 450, loss: 0.010965314693748951
step: 460, loss: 0.0001995009952224791
step: 470, loss: 6.0813654272351414e-05
step: 480, loss: 0.0001368995726807043
step: 490, loss: 0.0002580337750259787

```

step: 500, loss: 0.00478367367759347
step: 510, loss: 0.0008208160288631916
step: 520, loss: 0.004700292367488146
step: 530, loss: 0.009483062662184238
/usr/local/lib/python3.7/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=0.96040987424313, f1=0.9498829039812646,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006677740020677447
step: 10, loss: 0.0013716830871999264
step: 20, loss: 0.00021267200645525008
step: 30, loss: 0.00016507554391864687
step: 40, loss: 0.0007051859283819795
step: 50, loss: 0.0006206849357113242
step: 60, loss: 4.253070801496506e-05
step: 70, loss: 3.807448229053989e-05
step: 80, loss: 0.00044227924081496894
step: 90, loss: 0.007469952572137117
step: 100, loss: 0.0003671969461720437
step: 110, loss: 0.0028967128600925207
step: 120, loss: 8.930985495680943e-05
step: 130, loss: 0.001735667698085308
step: 140, loss: 0.00043663280666805804
step: 150, loss: 0.0005185853224247694
step: 160, loss: 0.0025510096456855536
step: 170, loss: 0.00015954078116919845
step: 180, loss: 0.06821132451295853
step: 190, loss: 0.002024831948801875
step: 200, loss: 0.0010982404928654432
step: 210, loss: 0.003074177773669362
step: 220, loss: 0.0012922529131174088
step: 230, loss: 0.006119093392044306
step: 240, loss: 0.0002646470966283232
step: 250, loss: 0.0019266940653324127
step: 260, loss: 0.00012865950702689588
step: 270, loss: 0.00011018146324204281
step: 280, loss: 0.011928590014576912
step: 290, loss: 0.0002340220526093617
step: 300, loss: 0.002736939350143075
step: 310, loss: 0.010999237187206745
step: 320, loss: 0.0008272352279163897
step: 330, loss: 0.0003809125046245754
step: 340, loss: 0.0003978977329097688
step: 350, loss: 0.005664281081408262

```

```

step: 360, loss: 0.0001313696993747726
step: 370, loss: 0.007718617096543312
step: 380, loss: 0.008573523722589016
step: 390, loss: 0.00017681217286735773
step: 400, loss: 0.00044397058081813157
step: 410, loss: 0.00027638114988803864
step: 420, loss: 0.00016714245430193841
step: 430, loss: 0.0033808329608291388
step: 440, loss: 0.0006887977360747755
step: 450, loss: 0.0013697363901883364
step: 460, loss: 0.0039059247355908155
step: 470, loss: 0.002223685849457979
step: 480, loss: 0.009741876274347305
step: 490, loss: 0.0011368915438652039
step: 500, loss: 0.024779340252280235
step: 510, loss: 4.138810618314892e-05
step: 520, loss: 0.00012198472541058436
step: 530, loss: 0.0007085914257913828
/usr/local/lib/python3.7/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=0.9576152771308802, f1=0.9471715755025713,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.015164010226727e-05
step: 10, loss: 0.00010282247967552394
step: 20, loss: 5.755138408858329e-05
step: 30, loss: 0.0001282298326259479
step: 40, loss: 3.512967668939382e-05
step: 50, loss: 2.8270018447074108e-05
step: 60, loss: 4.074254320585169e-05
step: 70, loss: 0.0002909085014835
step: 80, loss: 0.00011847104178741574
step: 90, loss: 0.0017788854893296957
step: 100, loss: 0.0010886102681979537
step: 110, loss: 3.186355752404779e-05
step: 120, loss: 0.0002705522929318249
step: 130, loss: 0.007617275696247816
step: 140, loss: 8.53521705721505e-05
step: 150, loss: 0.005197097547352314
step: 160, loss: 0.0031259541865438223
step: 170, loss: 1.6394531485275365e-05
step: 180, loss: 2.5083907530643046e-05
step: 190, loss: 0.0004959013895131648
step: 200, loss: 2.990569919347763e-05
step: 210, loss: 2.6339985197409987e-05

```



```

step: 220, loss: 0.002141796750947833
step: 230, loss: 1.311250980506884e-05
step: 240, loss: 0.0029518797527998686
step: 250, loss: 0.00017229949298780411
step: 260, loss: 0.002398652257397771
step: 270, loss: 0.008348818868398666
step: 280, loss: 0.00021534440747927874
step: 290, loss: 0.00011781955981859937
step: 300, loss: 1.9575798432924785e-05
step: 310, loss: 0.0008326011011376977
step: 320, loss: 0.00011085960431955755
step: 330, loss: 1.531431553303264e-05
step: 340, loss: 0.0007838673191145062
step: 350, loss: 2.3088176021701656e-05
step: 360, loss: 6.17432888248004e-05
step: 370, loss: 0.0014822069788351655
step: 380, loss: 0.00023977049568202347
step: 390, loss: 0.0011020069941878319
step: 400, loss: 0.00022496862220577896
step: 410, loss: 0.00301992311142385
step: 420, loss: 0.012672705575823784
step: 430, loss: 0.002388816559687257
step: 440, loss: 2.7636853701551445e-05
step: 450, loss: 0.0014289658283814788
step: 460, loss: 0.00013882384519092739
step: 470, loss: 0.0001521877129562199
step: 480, loss: 0.003995875827968121
step: 490, loss: 4.287801857572049e-05
step: 500, loss: 4.070257273269817e-05
step: 510, loss: 0.012651308439671993
step: 520, loss: 6.0736492741853e-05
step: 530, loss: 0.05280715227127075
/usr/local/lib/python3.7/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=0.9587389893370423, f1=0.9521597770552717,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.993093196186237e-05
step: 10, loss: 0.0011325327213853598
step: 20, loss: 0.0005160777946002781
step: 30, loss: 0.07115209847688675
step: 40, loss: 0.001489850226789713
step: 50, loss: 0.00013761347508989275
step: 60, loss: 0.0005342450458556414
step: 70, loss: 0.004024652298539877

```

step: 80, loss: 2.409069929854013e-05
step: 90, loss: 0.00014303455827757716
step: 100, loss: 0.005323102232068777
step: 110, loss: 0.0002647298970259726
step: 120, loss: 0.02505006082355976
step: 130, loss: 0.00025624301633797586
step: 140, loss: 3.673701576190069e-05
step: 150, loss: 0.00013659047544933856
step: 160, loss: 0.0004921836080029607
step: 170, loss: 1.667400101723615e-05
step: 180, loss: 1.5589957911288366e-05
step: 190, loss: 0.00021755688067059964
step: 200, loss: 0.002344646491110325
step: 210, loss: 2.4529830625397153e-05
step: 220, loss: 1.7027874491759576e-05
step: 230, loss: 1.594387140357867e-05
step: 240, loss: 1.0896235835389234e-05
step: 250, loss: 0.00048770959256216884
step: 260, loss: 0.0021502256859093904
step: 270, loss: 0.00010711993672885001
step: 280, loss: 2.0790317648788914e-05
step: 290, loss: 0.0030908691696822643
step: 300, loss: 0.0001864946389105171
step: 310, loss: 9.947687067324296e-05
step: 320, loss: 0.00023253924155142158
step: 330, loss: 0.003861349308863282
step: 340, loss: 1.6036803572205827e-05
step: 350, loss: 1.6829664673423395e-05
step: 360, loss: 5.500837505678646e-05
step: 370, loss: 1.5533138139289804e-05
step: 380, loss: 0.00019451735715847462
step: 390, loss: 0.020084800198674202
step: 400, loss: 3.716645733220503e-05
step: 410, loss: 0.0004722439043689519
step: 420, loss: 0.0013556020567193627
step: 430, loss: 0.0002832822210621089
step: 440, loss: 0.0015315655618906021
step: 450, loss: 3.7570913264062256e-05
step: 460, loss: 1.9794908439507708e-05
step: 470, loss: 0.005062208045274019
step: 480, loss: 0.00012045601761201397
step: 490, loss: 1.7303620552411303e-05
step: 500, loss: 2.9013184757786803e-05
step: 510, loss: 0.0029945774003863335
step: 520, loss: 0.007722585927695036
step: 530, loss: 0.00010844133794307709
/usr/local/lib/python3.7/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=0.9610027855153204, f1=0.9512308406874129,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.4502199519483838e-05
step: 10, loss: 1.5224937669700012e-05
step: 20, loss: 2.0853671230725013e-05
step: 30, loss: 1.2200000128359534e-05
step: 40, loss: 6.013646270730533e-05
step: 50, loss: 7.164808630477637e-05
step: 60, loss: 9.853239134827163e-06
step: 70, loss: 1.6349906218238175e-05
step: 80, loss: 0.0016129002906382084
step: 90, loss: 6.065180787118152e-05
step: 100, loss: 1.710218384687323e-05
step: 110, loss: 9.324265192844905e-06
step: 120, loss: 1.5206283023871947e-05
step: 130, loss: 0.0001793146220734343
step: 140, loss: 4.6631976147182286e-05
step: 150, loss: 0.00012275564949959517
step: 160, loss: 0.004038610029965639
step: 170, loss: 0.00016098501509986818
step: 180, loss: 2.288723408128135e-05
step: 190, loss: 7.400433241855353e-05
step: 200, loss: 2.2748083210899495e-05
step: 210, loss: 2.0506460714386776e-05
step: 220, loss: 8.547257311875e-05
step: 230, loss: 0.002223583636805415
step: 240, loss: 2.2365486074704677e-05
step: 250, loss: 0.00010063676018035039
step: 260, loss: 6.050614683772437e-05
step: 270, loss: 0.00018936203559860587
step: 280, loss: 1.749262082739733e-05
step: 290, loss: 9.044804755831137e-06
step: 300, loss: 2.475525252521038e-05
step: 310, loss: 1.3149966434866656e-05
step: 320, loss: 1.0661568012437783e-05
step: 330, loss: 1.3403417142399121e-05
step: 340, loss: 1.5601084669469856e-05
step: 350, loss: 6.560173915204359e-06
step: 360, loss: 0.10565929114818573
step: 370, loss: 1.3153427062206902e-05
step: 380, loss: 0.0014298813184723258
step: 390, loss: 3.585871309041977e-05
step: 400, loss: 2.200466587964911e-05
step: 410, loss: 0.00014135695528239012

```

```

step: 420, loss: 3.42440398526378e-05
step: 430, loss: 0.0002905924920924008
step: 440, loss: 0.00045256444718688726
step: 450, loss: 2.2554449969902635e-05
step: 460, loss: 7.698654371779412e-05
step: 470, loss: 0.003360514994710684
step: 480, loss: 7.0555743150180206e-06
step: 490, loss: 1.623044954612851e-05
step: 500, loss: 1.2315396816120483e-05
step: 510, loss: 1.75898676388897e-05
step: 520, loss: 1.618898386368528e-05
step: 530, loss: 1.960843837878201e-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=0.9594405594405595, f1=0.952513966480447,
best_f1=0.9529085872576176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011175140971317887
step: 10, loss: 7.635017391294241e-05
step: 20, loss: 2.70473537966609e-05
step: 30, loss: 5.7384397223358974e-05
step: 40, loss: 8.636554412078112e-05
step: 50, loss: 0.00010534541070228443
step: 60, loss: 4.1633240471128374e-05
step: 70, loss: 8.007725409697741e-05
step: 80, loss: 0.000522881222423166
step: 90, loss: 1.6047948520281352e-05
step: 100, loss: 1.9095212337560952e-05
step: 110, loss: 0.00011845504923257977
step: 120, loss: 2.7922938897972926e-05
step: 130, loss: 2.9452381568262354e-05
step: 140, loss: 7.247526809806004e-05
step: 150, loss: 0.0004035828751511872
step: 160, loss: 0.00034292470081709325
step: 170, loss: 7.732441736152396e-05
step: 180, loss: 2.7292780941934325e-05
step: 190, loss: 0.00020933093037456274
step: 200, loss: 5.9362933825468644e-05
step: 210, loss: 0.0018101050518453121
step: 220, loss: 1.1671093488985207e-05
step: 230, loss: 1.754548793542199e-05
step: 240, loss: 0.002075247699394822
step: 250, loss: 0.0010742578888311982
step: 260, loss: 1.605184661457315e-05
step: 270, loss: 0.10113199055194855

```

```

step: 280, loss: 2.662469159986358e-05
step: 290, loss: 8.07628111942904e-06
step: 300, loss: 0.00010686849418561906
step: 310, loss: 9.54530987655744e-05
step: 320, loss: 1.7485970602137968e-05
step: 330, loss: 0.0032432489097118378
step: 340, loss: 0.00011654620902845636
step: 350, loss: 1.2706712368526496e-05
step: 360, loss: 0.00019796693231910467
step: 370, loss: 0.0006376670207828283
step: 380, loss: 0.0019891427364200354
step: 390, loss: 0.00018147406808566302
step: 400, loss: 0.005026166792958975
step: 410, loss: 1.634161890251562e-05
step: 420, loss: 2.3903538021841086e-05
step: 430, loss: 0.000197634071810171
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 440, loss: 0.07277694344520569
step: 450, loss: 0.0008552305516786873
step: 460, loss: 0.02451540157198906
step: 470, loss: 1.9169761799275875e-05
step: 480, loss: 1.2196399438835215e-05
step: 490, loss: 0.003991311881691217
step: 500, loss: 3.4475371649023145e-05
step: 510, loss: 0.00417886208742857
step: 520, loss: 1.0225650839856826e-05
step: 530, loss: 1.7776614186004736e-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=0.9619666048237476, f1=0.954060324825986,
best_f1=0.954060324825986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.951546775875613e-05
step: 10, loss: 4.2986637708963826e-05
step: 20, loss: 0.003798002377152443
step: 30, loss: 0.004342358093708754
step: 40, loss: 1.4278771232056897e-05
step: 50, loss: 0.0005234070704318583
step: 60, loss: 0.0007313624955713749
step: 70, loss: 0.0001872810098575428
step: 80, loss: 2.5996694603236392e-05

```

```
step: 90, loss: 2.1097563148941845e-05
step: 100, loss: 5.666090146405622e-06
step: 110, loss: 0.00010079437925014645
step: 120, loss: 1.531759335193783e-05
step: 130, loss: 1.3749559911957476e-05
step: 140, loss: 1.5157359484874178e-05
step: 150, loss: 2.7850206606672145e-05
step: 160, loss: 1.671432619332336e-05
step: 170, loss: 2.271986159030348e-05
step: 180, loss: 1.7620124708628282e-05
step: 190, loss: 0.0040710968896746635
step: 200, loss: 0.0015193128492683172
step: 210, loss: 1.1324560546199791e-05
step: 220, loss: 0.00018305597768630832
step: 230, loss: 0.0001791054819477722
step: 240, loss: 0.000207171164220199
step: 250, loss: 0.016106605529785156
step: 260, loss: 1.263951799046481e-05
step: 270, loss: 1.7161646610475145e-05
step: 280, loss: 1.2587419405463152e-05
step: 290, loss: 0.000701303593814373
step: 300, loss: 1.8141356122214347e-05
step: 310, loss: 0.013555075041949749
step: 320, loss: 1.7977588868234307e-05
step: 330, loss: 0.0010466303210705519
step: 340, loss: 4.794459164259024e-05
step: 350, loss: 0.00036094640381634235
step: 360, loss: 1.3444283467833884e-05
step: 370, loss: 2.540534842410125e-05
step: 380, loss: 1.232302838616306e-05
step: 390, loss: 0.00033009867183864117
step: 400, loss: 0.016429070383310318
step: 410, loss: 0.00029665735200978816
step: 420, loss: 1.1700909453793429e-05
step: 430, loss: 7.562209702882683e-06
step: 440, loss: 4.0071226976579055e-05
step: 450, loss: 0.0014114441582933068
step: 460, loss: 0.000419285410316661
step: 470, loss: 1.0087912414746825e-05
step: 480, loss: 0.0012385235168039799
step: 490, loss: 0.0009309259476140141
step: 500, loss: 0.004446815699338913
step: 510, loss: 2.1753114197053947e-05
step: 520, loss: 1.72921281773597e-05
step: 530, loss: 9.495582162344363e-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=0.9614491407338597, f1=0.9541029207232267,
best_f1=0.954060324825986
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, 262.36it/s]
load_f1 = 0.9625866050808313
real_f1 = 0.9611829944547134
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, 202.21it/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 621 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: 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: srsly<1.1.0,>=0.1.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: plac<1.2.0,>=0.9.6 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.1.3)
Requirement already satisfied: 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: 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: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.6)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-

```



```

metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=d3c3bdfb454c23bdae8f5c6bffd3317e154f6348204292dac5ff27cf240fd070
  Stored in directory: /tmp/pip-ephem-wheel-cache-t6qgk8j9/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

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

```

```

--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm 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/

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

/usr/local/lib/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: 0.427725225687027
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=0.5128205128205129, f1=0.39215686274509803,
best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3762841522693634
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 2: dev_f1=0.34285714285714286, f1=0.3125, best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.46019336581230164
/usr/local/lib/python3.7/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=0.2947368421052632, f1=0.2978723404255319,
best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2926879823207855
/usr/local/lib/python3.7/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=0.36734693877551017, f1=0.36, best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31849992275238037
/usr/local/lib/python3.7/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=0.3636363636363636, f1=0.3561643835616438,
best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2797858715057373
/usr/local/lib/python3.7/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=0.35443037974683544, f1=0.3611111111111111,
best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5282694697380066
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
/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=0.4333333333333333, f1=0.36065573770491804,
best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5470772385597229
/usr/local/lib/python3.7/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=0.3373493975903615, f1=0.345679012345679,
best_f1=0.39215686274509803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33805471658706665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.8275862068965518, f1=0.7222222222222223,
best_f1=0.7222222222222223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2951990067958832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.8571428571428571, f1=0.8666666666666666,
best_f1=0.8666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18844695389270782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8750000000000001, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05596582219004631
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9032258064516129, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06951455026865005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019716570153832436
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013291634619235992
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8484848484848484
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 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, 120938.42it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.9285714285714286
real_f1 = 0.9285714285714286
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.86it/s]

```

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 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.")

step: 0, loss: 0.5659971833229065

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: 0.48051074147224426

step: 20, loss: 0.49816688895225525

step: 30, loss: 0.2612774074077606

step: 40, loss: 0.34072980284690857

step: 50, loss: 0.5141534805297852

step: 60, loss: 0.38033604621887207

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: 0.24901288747787476

step: 80, loss: 0.1804477423429489

step: 90, loss: 0.061728399246931076

step: 100, loss: 0.11146905273199081

step: 110, loss: 0.16420139372348785

step: 120, loss: 0.019341619685292244

step: 130, loss: 0.011702354997396469

step: 140, loss: 0.007815695367753506

step: 150, loss: 0.33396944403648376

step: 160, loss: 0.010326273739337921

step: 170, loss: 0.09169941395521164

step: 180, loss: 0.030014706775546074


```

step: 190, loss: 0.06353028863668442
step: 200, loss: 0.06469611823558807
step: 210, loss: 0.03160719573497772
step: 220, loss: 0.027492795139551163
step: 230, loss: 0.004777062684297562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.966740576496674, f1=0.9541899441340782,
best_f1=0.9541899441340782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01543548796325922
step: 10, loss: 0.0839538425207138
step: 20, loss: 0.019440783187747
step: 30, loss: 0.018295524641871452
step: 40, loss: 0.04861798509955406
step: 50, loss: 0.002891842508688569
step: 60, loss: 0.0018851226195693016
step: 70, loss: 0.006627082824707031
step: 80, loss: 0.006966578308492899
step: 90, loss: 0.012080123648047447
step: 100, loss: 0.008432727307081223
step: 110, loss: 0.004528888966888189
step: 120, loss: 0.0015580130275338888
step: 130, loss: 0.002073518233373761
step: 140, loss: 0.0031191501766443253
step: 150, loss: 0.11923276633024216
step: 160, loss: 0.004067402798682451
step: 170, loss: 0.0030309008434414864
step: 180, loss: 0.005389558617025614
step: 190, loss: 0.18277665972709656
step: 200, loss: 0.012488968670368195
step: 210, loss: 0.02150573395192623
step: 220, loss: 0.0007070833235047758
step: 230, loss: 0.0008198078721761703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9714285714285715, f1=0.9576174112256587,
best_f1=0.9576174112256587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.054978176951408386
step: 10, loss: 0.006036015693098307
step: 20, loss: 0.06892118602991104
step: 30, loss: 0.045790981501340866
step: 40, loss: 0.021370889618992805
step: 50, loss: 0.014046050608158112
step: 60, loss: 0.008814912289381027
step: 70, loss: 0.037086423486471176
step: 80, loss: 0.007370805833488703
step: 90, loss: 0.0037662507966160774
step: 100, loss: 0.0025520329363644123
step: 110, loss: 0.06421423703432083
step: 120, loss: 0.001424300018697977
step: 130, loss: 0.00650184229016304
step: 140, loss: 0.0007748364587314427
step: 150, loss: 0.05003233253955841
step: 160, loss: 0.004053290467709303
step: 170, loss: 0.0012304128613322973
step: 180, loss: 0.010751510970294476
step: 190, loss: 0.02010973170399666
step: 200, loss: 0.007553001865744591
step: 210, loss: 0.0018556989962235093
step: 220, loss: 0.06083577871322632
step: 230, loss: 0.03043394722044468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9852104664391355, f1=0.9817351598173515,
best_f1=0.9817351598173515
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011275085620582104
step: 10, loss: 0.0013331652153283358
step: 20, loss: 0.005336984526365995
step: 30, loss: 0.0009926040656864643
step: 40, loss: 0.06903916597366333
step: 50, loss: 0.0024223695509135723
step: 60, loss: 0.016539443284273148

```

```

step: 70, loss: 0.007735970430076122
step: 80, loss: 0.0034954186994582415
step: 90, loss: 0.0018370128236711025
step: 100, loss: 0.0014917552471160889
step: 110, loss: 0.0005536695243790746
step: 120, loss: 0.001202608342282474
step: 130, loss: 0.03961057588458061
step: 140, loss: 0.0020562694407999516
step: 150, loss: 0.0009245257824659348
step: 160, loss: 0.0019955276511609554
step: 170, loss: 0.005560228601098061
step: 180, loss: 0.02179901674389839
step: 190, loss: 0.000777982990257442
step: 200, loss: 0.008491797372698784
step: 210, loss: 0.0008896864019334316
step: 220, loss: 0.001986071467399597
step: 230, loss: 0.003908846061676741
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987598647125141, f1=0.984090909090909,
best_f1=0.984090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016232045600190759
step: 10, loss: 0.0014979176921769977
step: 20, loss: 0.000978988129645586
step: 30, loss: 0.0007078842609189451
step: 40, loss: 0.0009614262962713838
step: 50, loss: 0.001798517070710659
step: 60, loss: 0.0196093562990427
step: 70, loss: 0.010823789983987808
step: 80, loss: 0.01369659136980772
step: 90, loss: 0.013937296345829964
step: 100, loss: 0.0004556755593512207
step: 110, loss: 0.0005338857881724834
step: 120, loss: 0.0001435413141734898
step: 130, loss: 0.000261661276454106
step: 140, loss: 0.0015589639078825712
step: 150, loss: 0.16496047377586365
step: 160, loss: 0.0002446479629725218
step: 170, loss: 0.0017274919664487243
step: 180, loss: 0.001896550296805799

```

```

step: 190, loss: 0.010116537101566792
step: 200, loss: 0.000652291695587337
step: 210, loss: 0.0011707502417266369
step: 220, loss: 0.0009439076529815793
step: 230, loss: 0.0031078574247658253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898989898989898, f1=0.9898305084745763,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005992281367070973
step: 10, loss: 0.0014652629615738988
step: 20, loss: 0.0015739714726805687
step: 30, loss: 0.0008696025470271707
step: 40, loss: 0.0005285451770760119
step: 50, loss: 0.00046321528498083353
step: 60, loss: 0.0016198731027543545
step: 70, loss: 0.000360281381290406
step: 80, loss: 0.0008718567551113665
step: 90, loss: 0.0023604510352015495
step: 100, loss: 0.006273901090025902
step: 110, loss: 0.007696462795138359
step: 120, loss: 0.00026374487788416445
step: 130, loss: 0.0005006848950870335
step: 140, loss: 0.00034910987596958876
step: 150, loss: 0.00013951200526207685
step: 160, loss: 0.009879900142550468
step: 170, loss: 0.00015988406084943563
step: 180, loss: 0.000259108142927289
step: 190, loss: 0.0010203454876318574
step: 200, loss: 0.0011260370956733823
step: 210, loss: 0.00048621694440953434
step: 220, loss: 0.006334698759019375
step: 230, loss: 0.004997716750949621
/usr/local/lib/python3.7/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=0.983277591973244, f1=0.9776286353467561,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027615514118224382
step: 10, loss: 0.0021292974706739187
step: 20, loss: 0.0012269113212823868
step: 30, loss: 0.0010702407453209162
step: 40, loss: 0.001991966040804982
step: 50, loss: 0.000992869376204908
step: 60, loss: 0.0006096386932767928
step: 70, loss: 0.00044301999150775373
step: 80, loss: 0.0005591766675934196
step: 90, loss: 0.0014603780582547188
step: 100, loss: 0.0003786769520957023
step: 110, loss: 0.00029728325898759067
step: 120, loss: 0.0003652679151855409
step: 130, loss: 0.001927077304571867
step: 140, loss: 0.00021090314839966595
step: 150, loss: 0.030750615522265434
step: 160, loss: 0.00034411673550494015
step: 170, loss: 0.0002718174073379487
step: 180, loss: 0.0001897965557873249
step: 190, loss: 0.0001929308200487867
step: 200, loss: 0.00022481617634184659
step: 210, loss: 0.00031762017169967294
step: 220, loss: 0.0004683000151999295
step: 230, loss: 0.0006585437804460526
/usr/local/lib/python3.7/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=0.9865771812080537, f1=0.9887640449438202,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008992752991616726
step: 10, loss: 0.0031424658372998238
step: 20, loss: 0.003108066041022539
step: 30, loss: 0.001194942626170814
step: 40, loss: 0.0009286893182434142
step: 50, loss: 0.001665680785663426
step: 60, loss: 0.0005016261129640043
step: 70, loss: 0.00015129656821954995
step: 80, loss: 0.09591225534677505
step: 90, loss: 0.00170932337641716
step: 100, loss: 0.0011528786271810532
step: 110, loss: 0.0009238766506314278
step: 120, loss: 0.01585286296904087
step: 130, loss: 0.0022695071529597044
step: 140, loss: 0.0004018543695565313

```

```

step: 150, loss: 0.07255236059427261
step: 160, loss: 0.0009122495539486408
step: 170, loss: 0.0015904769534245133
step: 180, loss: 0.0004269997007213533
step: 190, loss: 0.0007733357488177717
step: 200, loss: 0.0013705651508644223
step: 210, loss: 0.0011762368958443403
step: 220, loss: 0.004371181596070528
step: 230, loss: 0.0007760628359392285
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9831271091113611,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023294783022720367
step: 10, loss: 0.0007411003462038934
step: 20, loss: 0.0005482141859829426
step: 30, loss: 0.0052682156674563885
step: 40, loss: 0.001106239971704781
step: 50, loss: 0.00238345586694777
step: 60, loss: 0.0006298673106357455
step: 70, loss: 0.05607059970498085
step: 80, loss: 0.0006757648661732674
step: 90, loss: 0.050847239792346954
step: 100, loss: 0.02663845382630825
step: 110, loss: 0.000167355960002169
step: 120, loss: 0.013104774057865143
step: 130, loss: 0.00046736231888644397
step: 140, loss: 0.00034761964343488216
step: 150, loss: 0.0007240533013828099
step: 160, loss: 0.0007940270006656647
step: 170, loss: 0.0002580849686637521
step: 180, loss: 0.012598160654306412
step: 190, loss: 0.000154605382704176
step: 200, loss: 0.00028572240262292325
step: 210, loss: 0.0008555843960493803
step: 220, loss: 0.0005452043260447681
step: 230, loss: 0.00025742745492607355
/usr/local/lib/python3.7/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=0.9778761061946903, f1=0.9865771812080537,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00047803213237784803
step: 10, loss: 0.0008547172183170915
step: 20, loss: 0.0010445110965520144
step: 30, loss: 0.00034975481685251
step: 40, loss: 0.007174962665885687
step: 50, loss: 0.0002316323370905593
step: 60, loss: 0.00040174208697862923
step: 70, loss: 0.04946171119809151
step: 80, loss: 0.0002876840007957071
step: 90, loss: 0.00041505767148919404
step: 100, loss: 0.00024208464310504496
step: 110, loss: 0.00028560811188071966
step: 120, loss: 0.0003219266945961863
step: 130, loss: 0.000386607222026214
step: 140, loss: 0.00033613157575018704
step: 150, loss: 0.04863540828227997
step: 160, loss: 0.0001843940990511328
step: 170, loss: 0.0003763343847822398
step: 180, loss: 0.0005262692575342953
step: 190, loss: 0.00039708384429104626
step: 200, loss: 0.0007937144837342203
step: 210, loss: 0.0010196997318416834
step: 220, loss: 0.03647390753030777
step: 230, loss: 0.0003412150254007429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921436588103255, f1=0.9843400447427293,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024650426348671317
step: 10, loss: 0.0029067823197692633
step: 20, loss: 0.0003090200771111995
step: 30, loss: 0.00019736887770704925
step: 40, loss: 3.69792542187497e-05
step: 50, loss: 7.666987221455202e-05
step: 60, loss: 0.015626076608896255
step: 70, loss: 0.00014454731717705727
step: 80, loss: 0.00029573068604804575
step: 90, loss: 0.04563587158918381
step: 100, loss: 0.0007268714252859354

```

```

step: 110, loss: 0.00022004333732184023
step: 120, loss: 0.00019494829757604748
step: 130, loss: 0.0004991085734218359
step: 140, loss: 0.005035163834691048
step: 150, loss: 0.00029983880813233554
step: 160, loss: 0.02826029062271118
step: 170, loss: 0.001930508529767394
step: 180, loss: 0.001451282761991024
step: 190, loss: 0.0009641917422413826
step: 200, loss: 0.01779255084693432
step: 210, loss: 0.00025041759363375604
step: 220, loss: 0.0003200116043444723
step: 230, loss: 0.00015629053814336658
/usr/local/lib/python3.7/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=0.9876543209876544,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036339517100714147
step: 10, loss: 0.00014177488628774881
step: 20, loss: 0.011314441449940205
step: 30, loss: 0.028317240998148918
step: 40, loss: 0.0009244014509022236
step: 50, loss: 0.0008782375371083617
step: 60, loss: 0.0004895787569694221
step: 70, loss: 0.0003135719452984631
step: 80, loss: 0.000164014330948703
step: 90, loss: 0.00043410283979028463
step: 100, loss: 0.0001992943143704906
step: 110, loss: 0.00014011339226271957
step: 120, loss: 0.00020170335483271629
step: 130, loss: 0.0001699544081930071
step: 140, loss: 0.00022749300114810467
step: 150, loss: 0.00024413748178631067
step: 160, loss: 0.00150096055585891
step: 170, loss: 0.0002352751325815916
step: 180, loss: 0.0001436940219718963
step: 190, loss: 0.0003955972788389772
step: 200, loss: 0.00011961481504840776
step: 210, loss: 0.0003148422692902386
step: 220, loss: 0.008271817117929459
step: 230, loss: 0.0005765573587268591
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.9843400447427293,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024680685601197183
step: 10, loss: 0.00034597038757056
step: 20, loss: 0.0005843776743859053
step: 30, loss: 0.0002510302874725312
step: 40, loss: 0.0009118468151427805
step: 50, loss: 0.002253307495266199
step: 60, loss: 0.0007751507218927145
step: 70, loss: 0.0005899373791180551
step: 80, loss: 0.00028039899189025164
step: 90, loss: 0.00043754593934863806
step: 100, loss: 0.0006978079909458756
step: 110, loss: 0.0005533922230824828
step: 120, loss: 0.0006744464626535773
step: 130, loss: 0.00028052108245901763
step: 140, loss: 0.00017578272672835737
step: 150, loss: 6.358505925163627e-05
step: 160, loss: 0.012609698809683323
step: 170, loss: 0.00017954390204977244
step: 180, loss: 0.016757264733314514
step: 190, loss: 0.00017953978385776281
step: 200, loss: 5.8393459767103195e-05
step: 210, loss: 0.0002576123515609652
step: 220, loss: 0.00013139043585397303
step: 230, loss: 0.0001634999644011259
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.9843400447427293,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012700428487733006
step: 10, loss: 0.00017813312297221273
step: 20, loss: 0.00020443153334781528
step: 30, loss: 0.00018988482770510018
step: 40, loss: 0.00014044017007108778
step: 50, loss: 0.00010977376223308966
step: 60, loss: 0.00012428818445187062
step: 70, loss: 0.00010864446085179225
step: 80, loss: 0.0001030442290357314
step: 90, loss: 0.0004644733853638172
step: 100, loss: 0.0001719228457659483

```

```

step: 110, loss: 0.0001974873011931777
step: 120, loss: 3.6266286770114675e-05
step: 130, loss: 0.0002365616528550163
step: 140, loss: 0.00013403651246335357
step: 150, loss: 0.00010479076445335522
step: 160, loss: 0.0002126401523128152
step: 170, loss: 0.00016011430125217885
step: 180, loss: 0.00010589040903141722
step: 190, loss: 0.00013526325346902013
step: 200, loss: 0.0002552873338572681
step: 210, loss: 0.00016681845590937883
step: 220, loss: 0.0002990749489981681
step: 230, loss: 0.00013027920795138925
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.9821029082774049,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011379738571122289
step: 10, loss: 0.0001755478442646563
step: 20, loss: 0.000308074289932847
step: 30, loss: 0.00019251568301115185
step: 40, loss: 8.312824502354488e-05
step: 50, loss: 0.00012708072608802468
step: 60, loss: 0.08016157895326614
step: 70, loss: 0.0005828819121234119
step: 80, loss: 0.00010837518493644893
step: 90, loss: 0.00011631500819930807
step: 100, loss: 0.0001360454480163753
step: 110, loss: 0.016173137351870537
step: 120, loss: 0.01809977926313877
step: 130, loss: 0.00010221003321930766
step: 140, loss: 0.011666535399854183
step: 150, loss: 0.00028312247013673186
step: 160, loss: 0.0190168134868145
step: 170, loss: 0.0001618016540305689
step: 180, loss: 0.00027569092344492674
step: 190, loss: 0.0001871929271146655
step: 200, loss: 0.0006312362384051085
step: 210, loss: 0.02160724252462387
step: 220, loss: 0.00018526441999711096
step: 230, loss: 0.0015723214019089937
/usr/local/lib/python3.7/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=0.9910514541387023, f1=0.984304932735426,
best_f1=0.9843400447427293
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
2473it [00:13, 182.75it/s]
load_f1 = 0.9909706546275394
real_f1 = 0.9920903954802259
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, 167.90it/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: 0.62354576587677
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.46135222911834717
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.4183085560798645
step: 30, loss: 0.34034034609794617
step: 40, loss: 0.24548153579235077
step: 50, loss: 0.1622491329908371
step: 60, loss: 0.2237061858177185
step: 70, loss: 0.3933168053627014
step: 80, loss: 0.20441380143165588
step: 90, loss: 0.12596578896045685
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.21908485889434814
step: 110, loss: 0.12918917834758759
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 120, loss: 0.12833552062511444
step: 130, loss: 0.06562091410160065
step: 140, loss: 0.14912879467010498
step: 150, loss: 0.06682717800140381
step: 160, loss: 0.07915650308132172
step: 170, loss: 0.11971594393253326
step: 180, loss: 0.03698517382144928
step: 190, loss: 0.10804463922977448
step: 200, loss: 0.05843871086835861
step: 210, loss: 0.07045862078666687
step: 220, loss: 0.005494887940585613
step: 230, loss: 0.17427803575992584
step: 240, loss: 0.12646622955799103
step: 250, loss: 0.025372754782438278
step: 260, loss: 0.16308781504631042
step: 270, loss: 0.36535075306892395
step: 280, loss: 0.03862139210104942
step: 290, loss: 0.08821168541908264
step: 300, loss: 0.03178178891539574
step: 310, loss: 0.08976603299379349
step: 320, loss: 0.030888812616467476
step: 330, loss: 0.128110870718956
step: 340, loss: 0.32706910371780396
step: 350, loss: 0.08875062316656113
step: 360, loss: 0.07333023846149445

```

```

step: 370, loss: 0.1561645269393921
step: 380, loss: 0.0742761567234993
step: 390, loss: 0.009654596447944641
step: 400, loss: 0.07948914915323257
step: 410, loss: 0.2522698640823364
step: 420, loss: 0.026555314660072327
step: 430, loss: 0.06356781721115112
step: 440, loss: 0.03549885004758835
step: 450, loss: 0.05733015015721321
step: 460, loss: 0.005328163038939238
step: 470, loss: 0.01240125298500061
step: 480, loss: 0.21921123564243317
step: 490, loss: 0.16984565556049347
step: 500, loss: 0.011622251011431217
step: 510, loss: 0.02181205153465271
step: 520, loss: 0.016637137159705162
step: 530, loss: 0.060058023780584335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.931098696461825, f1=0.9380776340110906,
best_f1=0.9380776340110906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06794765591621399
step: 10, loss: 0.03407455235719681
step: 20, loss: 0.003041462041437626
step: 30, loss: 0.03525930270552635
step: 40, loss: 0.04468545690178871
step: 50, loss: 0.02885611169040203
step: 60, loss: 0.010232294909656048
step: 70, loss: 0.003617505542933941
step: 80, loss: 0.06879764795303345
step: 90, loss: 0.02884780429303646
step: 100, loss: 0.25031623244285583
step: 110, loss: 0.007397225126624107
step: 120, loss: 0.09948895871639252
step: 130, loss: 0.002659651217982173
step: 140, loss: 0.04947986826300621
step: 150, loss: 0.02232300490140915
step: 160, loss: 0.029767058789730072
step: 170, loss: 0.0566880963742733
step: 180, loss: 0.03583696112036705

```

```

step: 190, loss: 0.021727753803133965
step: 200, loss: 0.21992236375808716
step: 210, loss: 0.038647133857011795
step: 220, loss: 0.004718360491096973
step: 230, loss: 0.03410141542553902
step: 240, loss: 0.07713805139064789
step: 250, loss: 0.014079012908041477
step: 260, loss: 0.028557628393173218
step: 270, loss: 0.01827867701649666
step: 280, loss: 0.03358590602874756
step: 290, loss: 0.025714486837387085
step: 300, loss: 0.05254777520895004
step: 310, loss: 0.030027642846107483
step: 320, loss: 0.06444284319877625
step: 330, loss: 0.0721956118941307
step: 340, loss: 0.07166546583175659
step: 350, loss: 0.0015683521050959826
step: 360, loss: 0.024173397570848465
step: 370, loss: 0.004867192357778549
step: 380, loss: 0.11360549181699753
step: 390, loss: 0.009405680000782013
step: 400, loss: 0.011143771931529045
step: 410, loss: 0.003792381612583995
step: 420, loss: 0.05973711609840393
step: 430, loss: 0.10459911078214645
step: 440, loss: 0.003902452066540718
step: 450, loss: 0.04106789827346802
step: 460, loss: 0.015404022298753262
step: 470, loss: 0.061787690967321396
step: 480, loss: 0.026836810633540154
step: 490, loss: 0.06269087642431259
step: 500, loss: 0.002135694259777665
step: 510, loss: 0.031386665999889374
step: 520, loss: 0.23787502944469452
step: 530, loss: 0.02975163422524929
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9525564256103178, f1=0.9497927222478121,
best_f1=0.9497927222478121
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12001524865627289

```

step: 10, loss: 0.11381173133850098
step: 20, loss: 0.004640913102775812
step: 30, loss: 0.041130051016807556
step: 40, loss: 0.03347093239426613
step: 50, loss: 0.010111640207469463
step: 60, loss: 0.018544606864452362
step: 70, loss: 0.013612248003482819
step: 80, loss: 0.15303510427474976
step: 90, loss: 0.006074517499655485
step: 100, loss: 0.01009821705520153
step: 110, loss: 0.012816179543733597
step: 120, loss: 0.15934951603412628
step: 130, loss: 0.04931383952498436
step: 140, loss: 0.010394726879894733
step: 150, loss: 0.03639513626694679
step: 160, loss: 0.013559320010244846
step: 170, loss: 0.0005125137977302074
step: 180, loss: 0.003913376946002245
step: 190, loss: 0.0013065349776297808
step: 200, loss: 0.022456634789705276
step: 210, loss: 0.052441179752349854
step: 220, loss: 0.046583905816078186
step: 230, loss: 0.016661588102579117
step: 240, loss: 0.05604318901896477
step: 250, loss: 0.029925227165222168
step: 260, loss: 0.06242147088050842
step: 270, loss: 0.000899051723536104
step: 280, loss: 0.007354370318353176
step: 290, loss: 0.0021498599089682102
step: 300, loss: 0.132982075214386
step: 310, loss: 0.06284776329994202
step: 320, loss: 0.05606665834784508
step: 330, loss: 0.004741129465401173
step: 340, loss: 0.0020864177495241165
step: 350, loss: 0.09129348397254944
step: 360, loss: 0.012684506364166737
step: 370, loss: 0.025213032960891724
step: 380, loss: 0.034704308956861496
step: 390, loss: 0.006338608451187611
step: 400, loss: 0.0520160011947155
step: 410, loss: 0.09010601043701172
step: 420, loss: 0.018382292240858078
step: 430, loss: 0.02257908135652542
step: 440, loss: 0.2242230474948883
step: 450, loss: 0.07210223376750946
step: 460, loss: 0.028941215947270393
step: 470, loss: 0.03935791924595833
step: 480, loss: 0.21847376227378845


```

step: 490, loss: 0.06195156276226044
step: 500, loss: 0.042766667902469635
step: 510, loss: 0.012769709341228008
step: 520, loss: 0.0011924096615985036
step: 530, loss: 0.0062119984067976475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9540176497909892, f1=0.9479553903345725,
best_f1=0.9479553903345725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015908364206552505
step: 10, loss: 0.003263202728703618
step: 20, loss: 0.014738059602677822
step: 30, loss: 0.13330569863319397
step: 40, loss: 0.037489499896764755
step: 50, loss: 0.00648427102714777
step: 60, loss: 0.004746764898300171
step: 70, loss: 0.0024157094303518534
step: 80, loss: 0.011588946916162968
step: 90, loss: 0.16068357229232788
step: 100, loss: 0.0031489720568060875
step: 110, loss: 0.04566476121544838
step: 120, loss: 0.0013440229231491685
step: 130, loss: 0.02060622163116932
step: 140, loss: 0.0168679840862751
step: 150, loss: 0.00262822094373405
step: 160, loss: 0.006271897349506617
step: 170, loss: 0.04076038673520088
step: 180, loss: 0.03778112307190895
step: 190, loss: 0.026609251275658607
step: 200, loss: 0.037717368453741074
step: 210, loss: 0.0005680415779352188
step: 220, loss: 0.015676984563469887
step: 230, loss: 0.011204191483557224
step: 240, loss: 0.04340367391705513
step: 250, loss: 0.1580767035484314
step: 260, loss: 0.011294765397906303
step: 270, loss: 0.19149601459503174
step: 280, loss: 0.009423298761248589
step: 290, loss: 0.06857451051473618
step: 300, loss: 0.000470963423140347

```

```

step: 310, loss: 0.0024684639647603035
step: 320, loss: 0.10148344933986664
step: 330, loss: 0.0069955564104020596
step: 340, loss: 0.0017031474271789193
step: 350, loss: 0.25823974609375
step: 360, loss: 0.04337712377309799
step: 370, loss: 0.0059172604233026505
step: 380, loss: 0.02794509381055832
step: 390, loss: 0.008069397881627083
step: 400, loss: 0.0011712569976225495
step: 410, loss: 0.00039856493822298944
step: 420, loss: 0.033240724354982376
step: 430, loss: 0.018428867682814598
step: 440, loss: 0.0012873488012701273
step: 450, loss: 0.011523791588842869
step: 460, loss: 0.024463023990392685
step: 470, loss: 0.0007971863378770649
step: 480, loss: 0.0026385884266346693
step: 490, loss: 0.018370384350419044
step: 500, loss: 0.03753020241856575
step: 510, loss: 0.029713446274399757
step: 520, loss: 0.0038966271094977856
step: 530, loss: 0.1316392719745636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9580838323353295, f1=0.9526436781609195,
best_f1=0.9526436781609195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009733598562888801
step: 10, loss: 0.03795653209090233
step: 20, loss: 0.016194092109799385
step: 30, loss: 0.003144233487546444
step: 40, loss: 0.0024417033419013023
step: 50, loss: 0.062043651938438416
step: 60, loss: 0.006437151692807674
step: 70, loss: 0.007672180887311697
step: 80, loss: 0.0019196951761841774
step: 90, loss: 0.029210487380623817
step: 100, loss: 0.09507037699222565
step: 110, loss: 0.0027659651823341846
step: 120, loss: 0.005004962906241417

```

```

step: 130, loss: 0.0337119996547699
step: 140, loss: 0.012625050731003284
step: 150, loss: 0.016064321622252464
step: 160, loss: 0.006502336356788874
step: 170, loss: 0.1131318211555481
step: 180, loss: 0.008252523839473724
step: 190, loss: 0.004675390664488077
step: 200, loss: 0.008993255905807018
step: 210, loss: 0.0017148390179499984
step: 220, loss: 0.009806477464735508
step: 230, loss: 0.0010864082723855972
step: 240, loss: 0.00226613599807024
step: 250, loss: 0.12478655576705933
step: 260, loss: 0.00010654965444700792
step: 270, loss: 0.05035748332738876
step: 280, loss: 0.006050311494618654
step: 290, loss: 0.014580406248569489
step: 300, loss: 0.027023635804653168
step: 310, loss: 0.035101987421512604
step: 320, loss: 0.03924413397908211
step: 330, loss: 0.012132403440773487
step: 340, loss: 0.010283579118549824
step: 350, loss: 7.024699880275875e-05
step: 360, loss: 0.00014334029401652515
step: 370, loss: 0.002135436749085784
step: 380, loss: 0.00014441921666730195
step: 390, loss: 0.00216672383248806
step: 400, loss: 0.009383638389408588
step: 410, loss: 0.10085504502058029
step: 420, loss: 0.17305514216423035
step: 430, loss: 0.001678081345744431
step: 440, loss: 0.0002442057302687317
step: 450, loss: 0.0016417847946286201
step: 460, loss: 0.0015286104753613472
step: 470, loss: 0.007661337964236736
step: 480, loss: 0.023410368710756302
step: 490, loss: 0.011749750934541225
step: 500, loss: 0.07961516082286835
step: 510, loss: 0.013914179988205433
step: 520, loss: 0.040588945150375366
step: 530, loss: 0.015671275556087494
/usr/local/lib/python3.7/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=0.9539748953974895, f1=0.9477611940298507,
best_f1=0.9526436781609195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023371567949652672
step: 10, loss: 0.00018580460164230317
step: 20, loss: 0.06407668441534042
step: 30, loss: 0.0038672317750751972
step: 40, loss: 0.00014283884956967086
step: 50, loss: 0.00031164733809418976
step: 60, loss: 0.08334187418222427
step: 70, loss: 0.012402240186929703
step: 80, loss: 0.0005363684613257647
step: 90, loss: 0.0004067292611580342
step: 100, loss: 0.00200855010189116
step: 110, loss: 0.0002667368098627776
step: 120, loss: 0.010040323249995708
step: 130, loss: 0.03617318347096443
step: 140, loss: 0.0020154155790805817
step: 150, loss: 0.0009200533968396485
step: 160, loss: 0.017861539497971535
step: 170, loss: 0.0004192400665488094
step: 180, loss: 0.003323999233543873
step: 190, loss: 0.01073844451457262
step: 200, loss: 0.03364607319235802
step: 210, loss: 0.0031698804814368486
step: 220, loss: 0.0032776901498436928
step: 230, loss: 0.23822470009326935
step: 240, loss: 0.017370549961924553
step: 250, loss: 0.015215473249554634
step: 260, loss: 0.008487009443342686
step: 270, loss: 0.0017382641090080142
step: 280, loss: 0.0031506225932389498
step: 290, loss: 0.0004039124178234488
step: 300, loss: 0.0023454031907022
step: 310, loss: 0.20205454528331757
step: 320, loss: 0.0003665265685413033
step: 330, loss: 0.013179480098187923
step: 340, loss: 0.0015434341039508581
step: 350, loss: 0.0029808059334754944
step: 360, loss: 0.0734112486243248
step: 370, loss: 0.001909574493765831
step: 380, loss: 0.0003716475039254874
step: 390, loss: 0.0002094469527946785
step: 400, loss: 0.0015575478319078684
step: 410, loss: 0.0003630876599345356
step: 420, loss: 0.01548149436712265
step: 430, loss: 0.001899780472740531
step: 440, loss: 0.006813570857048035
step: 450, loss: 0.2712818682193756
step: 460, loss: 0.0031885907519608736
```

```

step: 470, loss: 0.011108660139143467
step: 480, loss: 0.0036696908064186573
step: 490, loss: 0.0021530739031732082
step: 500, loss: 0.0006517653237096965
step: 510, loss: 0.11762271821498871
step: 520, loss: 0.0010694994125515223
step: 530, loss: 0.001665619551204145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.958100558659218, f1=0.9524697110904008,
best_f1=0.9524697110904008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019554710015654564
step: 10, loss: 0.015017981640994549
step: 20, loss: 0.0008929520845413208
step: 30, loss: 0.05492231249809265
step: 40, loss: 0.003519069869071245
step: 50, loss: 0.0019356731791049242
step: 60, loss: 0.005597498267889023
step: 70, loss: 0.0005954941734671593
step: 80, loss: 0.00045192011748440564
step: 90, loss: 0.00010352256504120305
step: 100, loss: 0.0009778795065358281
step: 110, loss: 0.0021554010454565287
step: 120, loss: 0.0005619609728455544
step: 130, loss: 0.00020430490258149803
step: 140, loss: 0.001371047692373395
step: 150, loss: 0.0018356121145188808
step: 160, loss: 0.010399003513157368
step: 170, loss: 0.0062721408903598785
step: 180, loss: 0.028933819383382797
step: 190, loss: 0.02213791199028492
step: 200, loss: 0.0004591250326484442
step: 210, loss: 0.0009986974764615297
step: 220, loss: 8.931468619266525e-05
step: 230, loss: 4.130265733692795e-05
step: 240, loss: 0.002695458009839058
step: 250, loss: 0.000547136936802417
step: 260, loss: 0.0010269159683957696
step: 270, loss: 0.00027956071426160634
step: 280, loss: 0.003189233597368002

```

```

step: 290, loss: 0.001112576574087143
step: 300, loss: 6.160579505376518e-05
step: 310, loss: 0.0009985853685066104
step: 320, loss: 0.0001910525606945157
step: 330, loss: 7.929593994049355e-05
step: 340, loss: 0.0007379701128229499
step: 350, loss: 0.002671394729986787
step: 360, loss: 0.001839721342548728
step: 370, loss: 0.03309716656804085
step: 380, loss: 0.017127826809883118
step: 390, loss: 0.023841310292482376
step: 400, loss: 0.005583236459642649
step: 410, loss: 0.01929575204849243
step: 420, loss: 0.003458854043856263
step: 430, loss: 0.0014746650122106075
step: 440, loss: 0.0010758761782199144
step: 450, loss: 0.002002120018005371
step: 460, loss: 0.00036545004695653915
step: 470, loss: 0.10957592725753784
step: 480, loss: 0.004757446702569723
step: 490, loss: 0.0008438574150204659
step: 500, loss: 0.0017605157336220145
step: 510, loss: 0.00049375876551494
step: 520, loss: 0.00048226831131614745
step: 530, loss: 0.0003590802079997957
/usr/local/lib/python3.7/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=0.9543761638733707, f1=0.954060324825986,
best_f1=0.9524697110904008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000401832367060706
step: 10, loss: 0.011425805278122425
step: 20, loss: 0.0004956299089826643
step: 30, loss: 7.231041672639549e-05
step: 40, loss: 3.0176877771737054e-05
step: 50, loss: 0.00012155016884207726
step: 60, loss: 6.861716246930882e-05
step: 70, loss: 0.00835429783910513
step: 80, loss: 0.0027022729627788067
step: 90, loss: 0.00174148625228554
step: 100, loss: 0.0016846583457663655
step: 110, loss: 0.00023879858781583607
step: 120, loss: 0.0014364030212163925
step: 130, loss: 0.00016223519924096763
step: 140, loss: 6.481411401182413e-05

```

```

step: 150, loss: 8.76146659720689e-05
step: 160, loss: 0.0088035361841321
step: 170, loss: 0.0685165598988533
step: 180, loss: 0.00011210848606424406
step: 190, loss: 0.0035987684968858957
step: 200, loss: 0.0007152238977141678
step: 210, loss: 0.07005436718463898
step: 220, loss: 0.0011932686902582645
step: 230, loss: 0.058267414569854736
step: 240, loss: 0.0018060986185446382
step: 250, loss: 0.00040932840784080327
step: 260, loss: 4.6453184040728956e-05
step: 270, loss: 0.0010120596271008253
step: 280, loss: 4.9432244850322604e-05
step: 290, loss: 0.01661088317632675
step: 300, loss: 0.0265260748565197
step: 310, loss: 0.002482965588569641
step: 320, loss: 0.02007010206580162
step: 330, loss: 0.0005675189313478768
step: 340, loss: 0.21241198480129242
step: 350, loss: 7.820456085028127e-05
step: 360, loss: 0.002619771286845207
step: 370, loss: 0.09391454607248306
step: 380, loss: 0.00027863020659424365
step: 390, loss: 0.002594087738543749
step: 400, loss: 0.0023547769524157047
step: 410, loss: 0.0015012937365099788
step: 420, loss: 0.0004875304875895381
step: 430, loss: 0.007807152345776558
step: 440, loss: 0.0006361746927723289
step: 450, loss: 0.0006548719247803092
step: 460, loss: 0.0076158307492733
step: 470, loss: 0.17314113676548004
step: 480, loss: 0.003995606210082769
step: 490, loss: 0.004540024790912867
step: 500, loss: 0.0018939288565889
step: 510, loss: 0.001644266420044005
step: 520, loss: 0.00010530635336181149
step: 530, loss: 4.013869693153538e-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 8: dev_f1=0.9584487534626038, f1=0.9570835256114444,
best_f1=0.9570835256114444

```

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

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

step: 0, loss: 2.920937549788505e-05

step: 10, loss: 0.008697797544300556

step: 20, loss: 0.00046775516238994896

step: 30, loss: 0.1429756134748459

step: 40, loss: 0.00173855887260288

step: 50, loss: 0.0017426378326490521

step: 60, loss: 0.0024245118256658316

step: 70, loss: 0.00041650794446468353

step: 80, loss: 0.0036118635907769203

step: 90, loss: 0.08479921519756317

step: 100, loss: 0.0009265088010579348

step: 110, loss: 0.0026520921383053064

step: 120, loss: 0.000442089163698256

step: 130, loss: 0.00014303362695500255

step: 140, loss: 0.00012238314957357943

step: 150, loss: 0.031870175153017044

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

step: 160, loss: 0.0029393816366791725

step: 170, loss: 0.0037182795349508524

step: 180, loss: 0.006014765705913305

step: 190, loss: 0.0006968588568270206

step: 200, loss: 0.0015005370369181037

step: 210, loss: 0.0019810334779322147

step: 220, loss: 0.00030585136846639216

step: 230, loss: 0.001488973619416356

step: 240, loss: 0.048082925379276276

step: 250, loss: 0.0020660259760916233

step: 260, loss: 0.00041010126005858183

step: 270, loss: 0.00039509820635430515

step: 280, loss: 0.02359536848962307

step: 290, loss: 4.246881871949881e-05

step: 300, loss: 7.361554889939725e-05

step: 310, loss: 0.17654436826705933

step: 320, loss: 0.008418474346399307

step: 330, loss: 0.0013387046055868268

step: 340, loss: 0.002970369765534997

step: 350, loss: 0.04466036707162857

step: 360, loss: 8.908970630727708e-05

step: 370, loss: 0.0002544310409575701

step: 380, loss: 0.002014370635151863

step: 390, loss: 0.009702923707664013

step: 400, loss: 0.0001574630878167227

step: 410, loss: 0.0846114456653595

step: 420, loss: 0.008293981663882732

step: 430, loss: 0.006528440862894058


```

step: 440, loss: 0.00017461099196225405
step: 450, loss: 0.0006916694110259414
step: 460, loss: 0.0010635638609528542
step: 470, loss: 0.00011968589387834072
step: 480, loss: 2.8002014005323872e-05
step: 490, loss: 0.0006292946054600179
step: 500, loss: 0.09821111708879471
step: 510, loss: 0.0012732547475025058
step: 520, loss: 0.023962317034602165
step: 530, loss: 0.04498862102627754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9597806215722121, f1=0.9487648673376029,
best_f1=0.9487648673376029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004327838309109211
step: 10, loss: 0.003217508550733328
step: 20, loss: 0.00011068099411204457
step: 30, loss: 0.0024706367403268814
step: 40, loss: 4.711453948402777e-05
step: 50, loss: 0.0019894663710147142
step: 60, loss: 3.720561653608456e-05
step: 70, loss: 0.005366274155676365
step: 80, loss: 2.562940790085122e-05
step: 90, loss: 5.1799099310301244e-05
step: 100, loss: 0.00045822703395970166
step: 110, loss: 0.0043738000094890594
step: 120, loss: 0.0004086799453943968
step: 130, loss: 0.012239035218954086
step: 140, loss: 0.0002321594365639612
step: 150, loss: 3.4230244637001306e-05
step: 160, loss: 0.0637894794344902
step: 170, loss: 6.014531754772179e-05
step: 180, loss: 0.0006283117108978331
step: 190, loss: 2.080156082229223e-05
step: 200, loss: 0.00872811209410429
step: 210, loss: 0.011459127068519592
step: 220, loss: 0.001373284263536334
step: 230, loss: 7.156385981943458e-05
step: 240, loss: 3.625718818511814e-05
step: 250, loss: 0.00016018129826989025

```

```

step: 260, loss: 2.4765236958046444e-05
step: 270, loss: 0.00010277390538249165
step: 280, loss: 0.016777649521827698
step: 290, loss: 0.0007936493493616581
step: 300, loss: 0.0005357988411560655
step: 310, loss: 0.04014508053660393
step: 320, loss: 0.00017647123604547232
step: 330, loss: 0.000956764561124146
step: 340, loss: 5.5471682571806014e-05
step: 350, loss: 0.00022748572519049048
step: 360, loss: 9.398077236255631e-05
step: 370, loss: 0.0009129103273153305
step: 380, loss: 0.005642361007630825
step: 390, loss: 4.797282599611208e-05
step: 400, loss: 8.954395161708817e-05
step: 410, loss: 0.0027445596642792225
step: 420, loss: 4.3424468458397314e-05
step: 430, loss: 3.839574856101535e-05
step: 440, loss: 3.7334091757657006e-05
step: 450, loss: 0.00039047878817655146
step: 460, loss: 3.724471389432438e-05
step: 470, loss: 0.001432871911674738
step: 480, loss: 3.668963836389594e-05
step: 490, loss: 3.4290675102965906e-05
step: 500, loss: 0.00986436940729618
step: 510, loss: 6.0444137488957494e-05
step: 520, loss: 3.559814649634063e-05
step: 530, loss: 0.004174166824668646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9635854341736695, f1=0.9565217391304348,
best_f1=0.9565217391304348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.337040470796637e-05
step: 10, loss: 0.0005596088012680411
step: 20, loss: 3.239059878978878e-05
step: 30, loss: 0.00021652124996762723
step: 40, loss: 7.279701094375923e-05
step: 50, loss: 3.006629594892729e-05
step: 60, loss: 3.204797758371569e-05
step: 70, loss: 3.703200491145253e-05

```

step: 80, loss: 0.029165947809815407
step: 90, loss: 0.007755569647997618
step: 100, loss: 3.094914791290648e-05
step: 110, loss: 2.779766873572953e-05
step: 120, loss: 2.787217454169877e-05
step: 130, loss: 2.616606798255816e-05
step: 140, loss: 0.0005311069544404745
step: 150, loss: 0.0032932674512267113
step: 160, loss: 0.013417760841548443
step: 170, loss: 0.00023673065879847854
step: 180, loss: 0.0010501507204025984
step: 190, loss: 0.0017882889369502664
step: 200, loss: 0.0003361113485880196
step: 210, loss: 0.0004710855719167739
step: 220, loss: 0.03132670000195503
step: 230, loss: 9.287631110055372e-05
step: 240, loss: 0.0028301100246608257
step: 250, loss: 0.00047113318578340113
step: 260, loss: 0.0042291060090065
step: 270, loss: 0.0025332325603812933
step: 280, loss: 0.055712468922138214
step: 290, loss: 0.00048657035222277045
step: 300, loss: 0.0001660243869991973
step: 310, loss: 0.0033227151725441217
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 320, loss: 0.000190965089132078
step: 330, loss: 5.8965495554730296e-05
step: 340, loss: 5.399657675297931e-05
step: 350, loss: 0.005567785352468491
step: 360, loss: 7.250110502354801e-05
step: 370, loss: 0.0038063835818320513
step: 380, loss: 0.000154260138515383
step: 390, loss: 0.001792891533114016
step: 400, loss: 0.00013982714153826237
step: 410, loss: 0.0010537832276895642
step: 420, loss: 0.00020468674483709037
step: 430, loss: 0.0073471288196742535
step: 440, loss: 0.0003707667929120362
step: 450, loss: 0.03791559860110283
step: 460, loss: 0.0006014127284288406
step: 470, loss: 0.03952144458889961
step: 480, loss: 0.0035621768329292536
step: 490, loss: 0.018110904842615128
step: 500, loss: 0.0013800989836454391
step: 510, loss: 0.00018039329734165221
step: 520, loss: 0.0022072328720241785
step: 530, loss: 0.004068322479724884
/usr/local/lib/python3.7/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=0.9587772116720705, f1=0.9510166358595193,
best_f1=0.9565217391304348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009203696390613914
step: 10, loss: 0.0027383940760046244
step: 20, loss: 0.005246992688626051
step: 30, loss: 0.0014582200674340129
step: 40, loss: 0.0005017311195842922
step: 50, loss: 7.584290869999677e-05
step: 60, loss: 0.000132103989017196
step: 70, loss: 0.00022759729472454637
step: 80, loss: 6.661663064733148e-05
step: 90, loss: 0.0007621551631018519
step: 100, loss: 0.09216675162315369
step: 110, loss: 0.00029164127772673965
step: 120, loss: 0.00043564094812609255
step: 130, loss: 0.0011534978402778506
step: 140, loss: 0.0022197149228304625
step: 150, loss: 0.00018136783910449594
step: 160, loss: 0.00529002770781517
step: 170, loss: 0.00017936425865627825
step: 180, loss: 7.630344043718651e-05
step: 190, loss: 0.0017116146627813578
step: 200, loss: 0.0014185821637511253
step: 210, loss: 0.0003017646668013185
step: 220, loss: 0.0002454248897265643
step: 230, loss: 4.828215969610028e-05
step: 240, loss: 0.0984722375869751
step: 250, loss: 0.00014824289246462286
step: 260, loss: 5.4886197176529095e-05
step: 270, loss: 0.0014149794587865472
step: 280, loss: 5.799889186164364e-05
step: 290, loss: 0.0011066681472584605
step: 300, loss: 0.0013374451082199812
step: 310, loss: 8.752772555453703e-05
step: 320, loss: 0.0004826991935260594
step: 330, loss: 0.0008801561198197305
step: 340, loss: 0.0001118179498007521
step: 350, loss: 0.0005720381159335375
step: 360, loss: 0.0001548401778563857
step: 370, loss: 0.02065812051296234
step: 380, loss: 0.0004126050916966051
step: 390, loss: 0.0014478567754849792
step: 400, loss: 5.844162660650909e-05

```

```

step: 410, loss: 0.0020204384345561266
step: 420, loss: 0.00012517556024249643
step: 430, loss: 0.0003734599449671805
step: 440, loss: 0.007463483139872551
step: 450, loss: 0.00014043276314623654
step: 460, loss: 0.00012635518214665353
step: 470, loss: 0.0037684720009565353
step: 480, loss: 0.0007239406113512814
step: 490, loss: 0.005924876779317856
step: 500, loss: 0.00036224848008714616
step: 510, loss: 0.004117561038583517
step: 520, loss: 0.00027849149773828685
step: 530, loss: 9.852238144958392e-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=0.9610027855153204, f1=0.9537465309898242,
best_f1=0.9565217391304348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.889051580103114e-05
step: 10, loss: 0.00014878633373882622
step: 20, loss: 7.683644798817113e-05
step: 30, loss: 0.00012227574188727885
step: 40, loss: 0.0007880459888838232
step: 50, loss: 0.0029521652031689882
step: 60, loss: 0.00046417114208452404
step: 70, loss: 0.0022268667817115784
step: 80, loss: 0.0007767474744468927
step: 90, loss: 0.00026523915585130453
step: 100, loss: 0.00016924944065976888
step: 110, loss: 0.00043619907228276134
step: 120, loss: 0.00025502449716441333
step: 130, loss: 6.968012894503772e-05
step: 140, loss: 0.0007685650489293039
step: 150, loss: 0.0009778562234714627
step: 160, loss: 0.0003038048162125051
step: 170, loss: 0.008589529432356358
step: 180, loss: 0.000360792939318344
step: 190, loss: 0.007102719973772764
step: 200, loss: 7.087624544510618e-05
step: 210, loss: 0.0003490410163067281
step: 220, loss: 6.433166709030047e-05
step: 230, loss: 0.0008759634220041335
step: 240, loss: 0.0009289688896387815
step: 250, loss: 0.0003410559438634664
step: 260, loss: 5.572381996898912e-05

```

```

step: 270, loss: 0.018764160573482513
step: 280, loss: 4.88249133923091e-05
step: 290, loss: 3.7099398468853906e-05
step: 300, loss: 0.00011903759150300175
step: 310, loss: 6.825307355029508e-05
step: 320, loss: 0.00011507282761158422
step: 330, loss: 0.0006829085177741945
step: 340, loss: 0.00018181004270445555
step: 350, loss: 3.43987085216213e-05
step: 360, loss: 0.10311373323202133
step: 370, loss: 6.353748176479712e-05
step: 380, loss: 0.00017562512948643416
step: 390, loss: 4.866480594500899e-05
step: 400, loss: 7.409972749883309e-05
step: 410, loss: 4.469036139198579e-05
step: 420, loss: 4.590874959831126e-05
step: 430, loss: 0.00014549907064065337
step: 440, loss: 0.00023696880089119077
step: 450, loss: 5.210201925365254e-05
step: 460, loss: 6.425563333323225e-05
step: 470, loss: 0.0010700048878788948
step: 480, loss: 3.0319668439915404e-05
step: 490, loss: 5.438004882307723e-05
step: 500, loss: 2.915740515163634e-05
step: 510, loss: 3.454374018474482e-05
step: 520, loss: 3.695759369293228e-05
step: 530, loss: 8.05590971140191e-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=0.9607390300230946, f1=0.9514563106796116,
best_f1=0.9565217391304348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012050550431013107
step: 10, loss: 0.0012032806407660246
step: 20, loss: 0.00029836565954610705
step: 30, loss: 0.00018705922411754727
step: 40, loss: 0.00012219318887218833
step: 50, loss: 8.300004265038297e-05
step: 60, loss: 3.865246253553778e-05
step: 70, loss: 0.0023573320358991623
step: 80, loss: 3.89527267543599e-05
step: 90, loss: 3.0889528716215864e-05
step: 100, loss: 0.0019597262144088745
step: 110, loss: 9.406442404724658e-05
step: 120, loss: 3.072933031944558e-05

```

```

step: 130, loss: 0.00018695862672757357
step: 140, loss: 0.0013618088560178876
step: 150, loss: 0.0022836211137473583
step: 160, loss: 3.9959457353688776e-05
step: 170, loss: 0.009285377338528633
step: 180, loss: 4.482131771510467e-05
step: 190, loss: 0.00010634153295541182
step: 200, loss: 3.236074553569779e-05
step: 210, loss: 0.0005400013760663569
step: 220, loss: 2.7693327865563333e-05
step: 230, loss: 2.9931750759715214e-05
step: 240, loss: 0.0008507549646310508
step: 250, loss: 0.0011860579252243042
step: 260, loss: 2.7015392333851196e-05
step: 270, loss: 0.0008903678972274065
step: 280, loss: 0.0003667299170047045
step: 290, loss: 2.361432598263491e-05
step: 300, loss: 3.304213169030845e-05
step: 310, loss: 3.6404588172445074e-05
step: 320, loss: 2.7510572181199677e-05
step: 330, loss: 0.0016115718754008412
step: 340, loss: 0.001015259069390595
step: 350, loss: 2.396448326180689e-05
step: 360, loss: 0.0001037034671753645
step: 370, loss: 0.00011131405335618183
step: 380, loss: 0.0016346893971785903
step: 390, loss: 3.742821718333289e-05
step: 400, loss: 0.0016536718467250466
step: 410, loss: 2.5592302336008288e-05
step: 420, loss: 2.5119294150499627e-05
step: 430, loss: 6.461524753831327e-05
step: 440, loss: 2.5789817300392315e-05
step: 450, loss: 2.315981691936031e-05
step: 460, loss: 0.02109498158097267
step: 470, loss: 2.1829953766427934e-05
step: 480, loss: 2.8300000849412754e-05
step: 490, loss: 0.0011766826501116157
step: 500, loss: 0.0023857136256992817
step: 510, loss: 0.009220881387591362
step: 520, loss: 0.000949014734942466
step: 530, loss: 0.002160429023206234
/usr/local/lib/python3.7/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=0.9602587800369685, f1=0.9509713228492137,
best_f1=0.9565217391304348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3703642000327818e-05
step: 10, loss: 0.001969036879017949
step: 20, loss: 0.0015456195687875152
step: 30, loss: 0.009586410596966743
step: 40, loss: 2.4761569875408895e-05
step: 50, loss: 0.0010521893855184317
step: 60, loss: 6.095445860410109e-05
step: 70, loss: 3.288110383437015e-05
step: 80, loss: 2.0328692698967643e-05
step: 90, loss: 5.523509025806561e-05
step: 100, loss: 2.121156830980908e-05
step: 110, loss: 3.22440464515239e-05
step: 120, loss: 2.2079524569562636e-05
step: 130, loss: 2.2548894776264206e-05
step: 140, loss: 2.285797381773591e-05
step: 150, loss: 2.26755491894437e-05
step: 160, loss: 2.4251241484307684e-05
step: 170, loss: 7.36949805286713e-05
step: 180, loss: 3.024818579433486e-05
step: 190, loss: 0.001838602009229362
step: 200, loss: 0.0009574202704243362
step: 210, loss: 2.673919516382739e-05
step: 220, loss: 2.229558231192641e-05
step: 230, loss: 2.21540431084577e-05
step: 240, loss: 2.6180858185398392e-05
step: 250, loss: 3.856847615679726e-05
step: 260, loss: 2.5011166144395247e-05
step: 270, loss: 2.1740544980275445e-05
step: 280, loss: 2.6091500330949202e-05
step: 290, loss: 0.0007786402711644769
step: 300, loss: 2.4210319679696113e-05
step: 310, loss: 0.03291601315140724
step: 320, loss: 0.007302509155124426
step: 330, loss: 2.2493035430670716e-05
step: 340, loss: 2.985386163345538e-05
step: 350, loss: 0.0014095032820478082
step: 360, loss: 3.348100290168077e-05
step: 370, loss: 2.484729884599801e-05
step: 380, loss: 2.3744691134197637e-05
step: 390, loss: 0.01756882667541504
step: 400, loss: 0.012755397707223892
step: 410, loss: 2.442261393298395e-05
step: 420, loss: 2.95998906949535e-05
step: 430, loss: 2.1323325199773535e-05
step: 440, loss: 2.5044739231816493e-05
step: 450, loss: 0.003478352213278413
step: 460, loss: 0.0014697008300572634
```



```

step: 470, loss: 2.2984750103205442e-05
step: 480, loss: 0.002185661345720291
step: 490, loss: 4.2778003262355924e-05
step: 500, loss: 0.002384177641943097
step: 510, loss: 3.060249218833633e-05
step: 520, loss: 2.2899046598467976e-05
step: 530, loss: 2.2843190890853293e-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=0.9615562760537285, f1=0.9530887134231305,
best_f1=0.9565217391304348
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 [00:25, 222.12it/s]
load_f1 = 0.9627213420316868
real_f1 = 0.9622025198320111
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, 175.15it/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: 0.5039631724357605
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.38115280866622925
step: 20, loss: 0.42228442430496216
step: 30, loss: 0.31102341413497925
step: 40, loss: 0.3703800439834595
step: 50, loss: 0.4318464696407318
step: 60, loss: 0.4896378815174103
step: 70, loss: 0.2961239516735077
step: 80, loss: 0.36560162901878357
step: 90, loss: 0.25633734464645386
step: 100, loss: 0.29627662897109985
step: 110, loss: 0.2623257637023926
step: 120, loss: 0.535537600517273
step: 130, loss: 0.2420603334903717
step: 140, loss: 0.3924984335899353
step: 150, loss: 0.3053397834300995
step: 160, loss: 0.28657805919647217
step: 170, loss: 0.21431589126586914
step: 180, loss: 0.21240508556365967
step: 190, loss: 0.40653830766677856
step: 200, loss: 0.14476704597473145

```

```

step: 210, loss: 0.20648127794265747
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5516074450084603, f1=0.6018018018018019,
best_f1=0.6018018018018019
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19531984627246857
step: 10, loss: 0.06832775473594666
step: 20, loss: 0.3374932110309601
step: 30, loss: 0.18387360870838165
step: 40, loss: 0.4782450497150421
step: 50, loss: 0.18848636746406555
step: 60, loss: 0.1948605626821518
step: 70, loss: 0.2259768396615982
step: 80, loss: 0.18246807157993317
step: 90, loss: 0.21047063171863556
step: 100, loss: 0.3737275004386902
step: 110, loss: 0.2976541519165039
step: 120, loss: 0.07603881508111954
step: 130, loss: 0.13933716714382172
step: 140, loss: 0.136368066072464
step: 150, loss: 0.2408292442560196
step: 160, loss: 0.033914752304553986
step: 170, loss: 0.4283508360385895
step: 180, loss: 0.2234189510345459
step: 190, loss: 0.2228282243013382
step: 200, loss: 0.15325114130973816
step: 210, loss: 0.18889875710010529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6504672897196262, f1=0.7001795332136446,
best_f1=0.7001795332136446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07689227163791656

```

```

step: 10, loss: 0.02201790176331997
step: 20, loss: 0.16039632260799408
step: 30, loss: 0.025672556832432747
step: 40, loss: 0.29913756251335144
step: 50, loss: 0.09744185954332352
step: 60, loss: 0.16920162737369537
step: 70, loss: 0.09686242043972015
step: 80, loss: 0.08690142631530762
step: 90, loss: 0.03894360736012459
step: 100, loss: 0.11280211061239243
step: 110, loss: 0.10804009437561035
step: 120, loss: 0.1558312326669693
step: 130, loss: 0.11076682806015015
step: 140, loss: 0.12153411656618118
step: 150, loss: 0.12011868506669998
step: 160, loss: 0.15600618720054626
step: 170, loss: 0.13545061647891998
step: 180, loss: 0.10411428660154343
step: 190, loss: 0.023338573053479195
step: 200, loss: 0.04603707790374756
step: 210, loss: 0.15955978631973267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.6811023622047243, f1=0.688976377952756,
best_f1=0.688976377952756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04866918921470642
step: 10, loss: 0.07656870037317276
step: 20, loss: 0.06710298359394073
step: 30, loss: 0.09997434914112091
step: 40, loss: 0.025305721908807755
step: 50, loss: 0.04492538049817085
step: 60, loss: 0.19936586916446686
step: 70, loss: 0.07610124349594116
step: 80, loss: 0.04598259925842285
step: 90, loss: 0.04006896913051605
step: 100, loss: 0.14692822098731995
step: 110, loss: 0.2060239017009735
step: 120, loss: 0.09825494885444641
step: 130, loss: 0.19011493027210236
step: 140, loss: 0.12080913782119751

```

```

step: 150, loss: 0.0761934369802475
step: 160, loss: 0.08081161230802536
step: 170, loss: 0.16788361966609955
step: 180, loss: 0.08184593170881271
step: 190, loss: 0.11451543122529984
step: 200, loss: 0.04656157270073891
step: 210, loss: 0.23772135376930237
/usr/local/lib/python3.7/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=0.6748971193415637, f1=0.7242798353909465,
best_f1=0.688976377952756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12421674281358719
step: 10, loss: 0.04790722206234932
step: 20, loss: 0.05782361701130867
step: 30, loss: 0.008396651595830917
step: 40, loss: 0.027555691078305244
step: 50, loss: 0.07737985253334045
step: 60, loss: 0.0337803028523922
step: 70, loss: 0.17636334896087646
step: 80, loss: 0.04550732672214508
step: 90, loss: 0.06569764763116837
step: 100, loss: 0.05916360020637512
step: 110, loss: 0.08809687197208405
step: 120, loss: 0.07735300064086914
step: 130, loss: 0.04123096913099289
step: 140, loss: 0.07960819453001022
step: 150, loss: 0.024122267961502075
step: 160, loss: 0.02727452479302883
step: 170, loss: 0.026468489319086075
step: 180, loss: 0.07187477499246597
step: 190, loss: 0.09326522052288055
step: 200, loss: 0.019288720563054085
step: 210, loss: 0.059509746730327606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6907216494845361, f1=0.690909090909091,
best_f1=0.690909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015058434568345547
step: 10, loss: 0.040111932903528214
step: 20, loss: 0.01595287024974823
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.015015187673270702
step: 40, loss: 0.024112723767757416
step: 50, loss: 0.021092934533953667
step: 60, loss: 0.13142940402030945
step: 70, loss: 0.06422074884176254
step: 80, loss: 0.1095643863081932
step: 90, loss: 0.03628576919436455
step: 100, loss: 0.13761885464191437
step: 110, loss: 0.05850543826818466
step: 120, loss: 0.01851501874625683
step: 130, loss: 0.0141630033031106
step: 140, loss: 0.11874278634786606
step: 150, loss: 0.031051497906446457
step: 160, loss: 0.02836744114756584
step: 170, loss: 0.04864783212542534
step: 180, loss: 0.017015252262353897
step: 190, loss: 0.08665280789136887
step: 200, loss: 0.04008936136960983
step: 210, loss: 0.023412901908159256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6924643584521385, f1=0.7313131313131312,
best_f1=0.7313131313131312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08075583726167679
step: 10, loss: 0.0038412827998399734
step: 20, loss: 0.010544821619987488
step: 30, loss: 0.005085440818220377
step: 40, loss: 0.01627250574529171
step: 50, loss: 0.10977397114038467
step: 60, loss: 0.08827152103185654
step: 70, loss: 0.04912440478801727
step: 80, loss: 0.04722290858626366
step: 90, loss: 0.2913053333759308
step: 100, loss: 0.020664812996983528
step: 110, loss: 0.036126431077718735

```

```

step: 120, loss: 0.02794388309121132
step: 130, loss: 0.045707255601882935
step: 140, loss: 0.05111542344093323
step: 150, loss: 0.050099339336156845
step: 160, loss: 0.2587006390094757
step: 170, loss: 0.02038581110537052
step: 180, loss: 0.013765977695584297
step: 190, loss: 0.012558807618916035
step: 200, loss: 0.03033345565199852
step: 210, loss: 0.21387623250484467
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6998011928429424, f1=0.7201565557729942,
best_f1=0.7201565557729942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06564673781394958
step: 10, loss: 0.003729618387296796
step: 20, loss: 0.03547164425253868
step: 30, loss: 0.05387784540653229
step: 40, loss: 0.02556012198328972
step: 50, loss: 0.0010385505156591535
step: 60, loss: 0.0038266717456281185
step: 70, loss: 0.15231311321258545
step: 80, loss: 0.06472976505756378
step: 90, loss: 0.16693484783172607
step: 100, loss: 0.040724631398916245
step: 110, loss: 0.005142070818692446
step: 120, loss: 0.06871714442968369
step: 130, loss: 0.007398610003292561
step: 140, loss: 0.06472066044807434
step: 150, loss: 0.05438714101910591
step: 160, loss: 0.044268134981393814
step: 170, loss: 0.09747985005378723
step: 180, loss: 0.025866307318210602
step: 190, loss: 0.009675244800746441
step: 200, loss: 0.010072408244013786
step: 210, loss: 0.10171365737915039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```


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=0.7051792828685258, f1=0.6982248520710059,

best_f1=0.6982248520710059

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

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

step: 0, loss: 0.034733932465314865

step: 10, loss: 0.04505675658583641

step: 20, loss: 0.013144725002348423

step: 30, loss: 0.001886689686216414

step: 40, loss: 0.026585472747683525

step: 50, loss: 0.006004793103784323

step: 60, loss: 0.00024031060456763953

step: 70, loss: 0.021660154685378075

step: 80, loss: 0.0020440651569515467

step: 90, loss: 0.004657357465475798

step: 100, loss: 0.02042985148727894

step: 110, loss: 0.037220269441604614

step: 120, loss: 0.059794530272483826

step: 130, loss: 0.04413097724318504

step: 140, loss: 0.14686548709869385

step: 150, loss: 0.07425201684236526

step: 160, loss: 0.0622386671602726

step: 170, loss: 0.08831024914979935

step: 180, loss: 0.0675010085105896

step: 190, loss: 0.005657005589455366

step: 200, loss: 0.0020139736589044333

step: 210, loss: 0.009761122055351734

/usr/local/lib/python3.7/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=0.7034764826175869, f1=0.7087576374745417,

best_f1=0.6982248520710059

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

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

step: 0, loss: 0.0013776790583506227

step: 10, loss: 0.0010415567085146904

step: 20, loss: 0.009586029686033726

step: 30, loss: 0.06830163300037384

step: 40, loss: 0.004011033568531275

step: 50, loss: 0.013378525152802467

step: 60, loss: 0.0002710068947635591

step: 70, loss: 0.010444943793118

step: 80, loss: 0.013459723442792892

step: 90, loss: 0.0102797020226717

```

step: 100, loss: 0.0401233546435833
step: 110, loss: 0.000948433531448245
step: 120, loss: 0.0193052738904953
step: 130, loss: 0.025944538414478302
step: 140, loss: 0.012318538501858711
step: 150, loss: 0.011184030212461948
step: 160, loss: 0.018209759145975113
step: 170, loss: 0.013073659501969814
step: 180, loss: 0.23543071746826172
step: 190, loss: 0.05037393793463707
step: 200, loss: 0.01479083951562643
step: 210, loss: 0.028028180822730064
/usr/local/lib/python3.7/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=0.701627486437613, f1=0.7160940325497287,
best_f1=0.6982248520710059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022193601354956627
step: 10, loss: 0.0015037215780466795
step: 20, loss: 0.03210267797112465
step: 30, loss: 0.03628208860754967
step: 40, loss: 0.008332393132150173
step: 50, loss: 0.005386761389672756
step: 60, loss: 0.00866628997027874
step: 70, loss: 0.011967970058321953
step: 80, loss: 0.02794921025633812
step: 90, loss: 0.06175227090716362
step: 100, loss: 0.0018800696125254035
step: 110, loss: 0.0006925025954842567
step: 120, loss: 0.014811167493462563
step: 130, loss: 0.005720621906220913
step: 140, loss: 0.04348552227020264
step: 150, loss: 0.012913430109620094
step: 160, loss: 0.0020943081472069025
step: 170, loss: 0.033977214246988297
step: 180, loss: 0.003830925328657031
step: 190, loss: 0.09591741859912872
step: 200, loss: 0.0005984109593555331
step: 210, loss: 0.0005261664045974612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7054108216432866, f1=0.7056451612903225,
best_f1=0.7056451612903225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003792949428316206
step: 10, loss: 0.0009425542084500194
step: 20, loss: 0.008538256399333477
step: 30, loss: 0.004307609982788563
step: 40, loss: 0.0001627802412258461
step: 50, loss: 0.017717653885483742
step: 60, loss: 0.00024349484010599554
step: 70, loss: 0.010628712363541126
step: 80, loss: 0.026325955986976624
step: 90, loss: 0.0018808948807418346
step: 100, loss: 3.97031290049199e-05
step: 110, loss: 0.015296232886612415
step: 120, loss: 0.000164057215442881
step: 130, loss: 0.03309377282857895
step: 140, loss: 0.0071983328089118
step: 150, loss: 5.222877371124923e-05
step: 160, loss: 0.00014793105947319418
step: 170, loss: 0.04163498058915138
step: 180, loss: 0.018968509510159492
step: 190, loss: 0.0005533889052458107
step: 200, loss: 0.0030447582248598337
step: 210, loss: 0.00044556462671607733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7061143984220907, f1=0.7142857142857142,
best_f1=0.7142857142857142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025476790964603424
step: 10, loss: 0.03227686882019043
step: 20, loss: 0.00026032933965325356
step: 30, loss: 0.00028881747857667506
step: 40, loss: 0.038958024233579636
step: 50, loss: 0.07194583117961884
step: 60, loss: 0.00702873757109046
step: 70, loss: 0.0002881904656533152

```

```

step: 80, loss: 0.05560843273997307
step: 90, loss: 0.005118633154779673
step: 100, loss: 0.007874679751694202
step: 110, loss: 0.08359500020742416
step: 120, loss: 0.007334364112466574
step: 130, loss: 0.0004144787380937487
step: 140, loss: 0.020557522773742676
step: 150, loss: 0.007792005781084299
step: 160, loss: 0.015050729736685753
step: 170, loss: 0.0006371127674356103
step: 180, loss: 0.002461525844410062
step: 190, loss: 0.03222263604402542
step: 200, loss: 0.05656052753329277
step: 210, loss: 0.06896460801362991
/usr/local/lib/python3.7/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=0.7012448132780082,
best_f1=0.7142857142857142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006541885086335242
step: 10, loss: 0.02537345513701439
step: 20, loss: 0.0005773490411229432
step: 30, loss: 0.0006571328267455101
step: 40, loss: 0.010825617238879204
step: 50, loss: 0.0014058406231924891
step: 60, loss: 0.016114473342895508
step: 70, loss: 0.015752548351883888
step: 80, loss: 0.031759776175022125
step: 90, loss: 7.7520300692413e-05
step: 100, loss: 0.006883643101900816
step: 110, loss: 0.0038016457110643387
step: 120, loss: 0.00015926064224913716
step: 130, loss: 0.00021052056399639696
step: 140, loss: 0.0005102447466924787
step: 150, loss: 0.044783856719732285
step: 160, loss: 0.00014648355136159807
step: 170, loss: 0.024725129827857018
step: 180, loss: 0.00013072032015770674
step: 190, loss: 0.029830675572156906
step: 200, loss: 0.14131805300712585
step: 210, loss: 0.0014632996171712875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.7095435684647303, f1=0.7053941908713692,
best_f1=0.7053941908713692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020218871941324323
step: 10, loss: 0.0015063363825902343
step: 20, loss: 0.003088679164648056
step: 30, loss: 0.006963340565562248
step: 40, loss: 0.008670192211866379
step: 50, loss: 0.00015123444609344006
step: 60, loss: 0.014368684962391853
step: 70, loss: 0.008633589372038841
step: 80, loss: 0.05785684287548065
step: 90, loss: 0.0008859411464072764
step: 100, loss: 0.0015620364574715495
step: 110, loss: 0.017120549455285072
step: 120, loss: 0.0005968929035589099
step: 130, loss: 0.00022763133165426552
step: 140, loss: 0.035732634365558624
step: 150, loss: 0.0022431630641222
step: 160, loss: 0.014733393676578999
step: 170, loss: 0.05445590242743492
step: 180, loss: 0.00025291467318311334
step: 190, loss: 0.0013310828944668174
step: 200, loss: 0.0001259907439816743
step: 210, loss: 0.027231959626078606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7148760330578512, f1=0.709278350515464,
best_f1=0.709278350515464
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 [00:06, 365.25it/s]
load_f1 = 0.7137096774193549
real_f1 = 0.7032520325203253
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, 172.49it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.47667890787124634

step: 10, loss: 0.41507479548454285

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

step: 20, loss: 0.25855153799057007

step: 30, loss: 0.3966313302516937

step: 40, loss: 0.25252336263656616

step: 50, loss: 0.3077199459075928

step: 60, loss: 0.47388845682144165

step: 70, loss: 0.41108131408691406

step: 80, loss: 0.18205903470516205

step: 90, loss: 0.30844646692276

step: 100, loss: 0.3931911289691925

step: 110, loss: 0.25349846482276917

step: 120, loss: 0.30955585837364197

step: 130, loss: 0.32920828461647034

step: 140, loss: 0.19527967274188995

step: 150, loss: 0.3137543797492981

step: 160, loss: 0.24558351933956146

step: 170, loss: 0.3945341110229492

step: 180, loss: 0.15318681299686432

step: 190, loss: 0.15883903205394745

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

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

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

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: 0.3576466739177704

step: 10, loss: 0.35241052508354187

step: 20, loss: 0.6716004610061646

step: 30, loss: 0.2649361193180084

step: 40, loss: 0.6348498463630676

step: 50, loss: 0.31952789425849915

step: 60, loss: 0.4503556489944458

step: 70, loss: 0.2982625961303711

step: 80, loss: 0.14191116392612457


```

step: 90, loss: 0.28068849444389343
step: 100, loss: 0.2203872948884964
step: 110, loss: 0.3445436358451843
step: 120, loss: 0.19689014554023743
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.496900737285614
step: 140, loss: 0.28791818022727966
step: 150, loss: 0.3186170756816864
step: 160, loss: 0.2789548933506012
step: 170, loss: 0.2632408142089844
step: 180, loss: 0.1921081691980362
step: 190, loss: 0.24735204875469208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.26520681265206814, f1=0.28467153284671537,
best_f1=0.28467153284671537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3855230510234833
step: 10, loss: 0.3179030120372772
step: 20, loss: 0.4840421974658966
step: 30, loss: 0.2365088313817978
step: 40, loss: 0.1599818766117096
step: 50, loss: 0.34608688950538635
step: 60, loss: 0.1925150454044342
step: 70, loss: 0.3638212978839874
step: 80, loss: 0.27685412764549255
step: 90, loss: 0.2673623859882355
step: 100, loss: 0.30706560611724854
step: 110, loss: 0.5159093737602234
step: 120, loss: 0.38428324460983276
step: 130, loss: 0.10832131654024124
step: 140, loss: 0.10567661374807358
step: 150, loss: 0.287699431180954
step: 160, loss: 0.3844121992588043
step: 170, loss: 0.3349374532699585
step: 180, loss: 0.24367214739322662
step: 190, loss: 0.14877964556217194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.6571428571428571, f1=0.6153846153846154,

best_f1=0.6153846153846154

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

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

step: 0, loss: 0.07152842730283737

step: 10, loss: 0.09665963053703308

step: 20, loss: 0.16510802507400513

step: 30, loss: 0.1682802140712738

step: 40, loss: 0.22886012494564056

step: 50, loss: 0.10216837376356125

step: 60, loss: 0.24161745607852936

step: 70, loss: 0.07559863477945328

step: 80, loss: 0.06507985293865204

step: 90, loss: 0.03259756416082382

step: 100, loss: 0.2972187101840973

step: 110, loss: 0.28216299414634705

step: 120, loss: 0.08984801918268204

step: 130, loss: 0.11666533350944519

step: 140, loss: 0.17050445079803467

step: 150, loss: 0.08564560860395432

step: 160, loss: 0.15117588639259338

step: 170, loss: 0.10424375534057617

step: 180, loss: 0.13513119518756866

step: 190, loss: 0.2155521810054779

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

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

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

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=0.7654320987654321, f1=0.7345971563981042,

best_f1=0.7345971563981042

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

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

step: 0, loss: 0.0976204127073288

step: 10, loss: 0.08160708844661713

step: 20, loss: 0.036321010440588

step: 30, loss: 0.12311485409736633

step: 40, loss: 0.04764492064714432

step: 50, loss: 0.15288203954696655

step: 60, loss: 0.008159981109201908

step: 70, loss: 0.09741435945034027

```

step: 80, loss: 0.04005083814263344
step: 90, loss: 0.026100847870111465
step: 100, loss: 0.061566311866045
step: 110, loss: 0.11797399073839188
step: 120, loss: 0.05884180963039398
step: 130, loss: 0.2174742966890335
step: 140, loss: 0.06010868400335312
step: 150, loss: 0.07575654238462448
step: 160, loss: 0.05419732257723808
step: 170, loss: 0.06365139037370682
step: 180, loss: 0.023216519504785538
step: 190, loss: 0.02050526812672615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8047619047619048, f1=0.7437641723356008,
best_f1=0.7437641723356008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19737231731414795
step: 10, loss: 0.01734691858291626
step: 20, loss: 0.03492635861039162
step: 30, loss: 0.011355521157383919
step: 40, loss: 0.010702531784772873
step: 50, loss: 0.03417729586362839
step: 60, loss: 0.04149850085377693
step: 70, loss: 0.04453716054558754
step: 80, loss: 0.1232953891158104
step: 90, loss: 0.08811720460653305
step: 100, loss: 0.08947435766458511
step: 110, loss: 0.009258953854441643
step: 120, loss: 0.023699304088950157
step: 130, loss: 0.0244044978171587
step: 140, loss: 0.0018265609396621585
step: 150, loss: 0.030358918011188507
step: 160, loss: 0.2579175531864166
step: 170, loss: 0.23134228587150574
step: 180, loss: 0.011129306629300117
step: 190, loss: 0.038271281868219376
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.8155844155844156, f1=0.8153846153846155,

best_f1=0.8153846153846155

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

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

step: 0, loss: 0.05092046037316322

step: 10, loss: 0.011653375811874866

step: 20, loss: 0.0052711134776473045

step: 30, loss: 0.024329913780093193

step: 40, loss: 0.022454533725976944

step: 50, loss: 0.012242581695318222

step: 60, loss: 0.0480155423283577

step: 70, loss: 0.004571843892335892

step: 80, loss: 0.003107329597696662

step: 90, loss: 0.03848903998732567

step: 100, loss: 0.27608081698417664

step: 110, loss: 0.20071960985660553

step: 120, loss: 0.026422105729579926

step: 130, loss: 0.004510423634201288

step: 140, loss: 0.07118266075849533

step: 150, loss: 0.02890283428132534

step: 160, loss: 0.07366655766963959

step: 170, loss: 0.13061071932315826

step: 180, loss: 0.009501403197646141

step: 190, loss: 0.022600188851356506

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

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

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

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=0.8186528497409327, f1=0.8030303030303031,

best_f1=0.8030303030303031

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

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

step: 0, loss: 0.018466230481863022

step: 10, loss: 0.02259521558880806

step: 20, loss: 0.0022683285642415285

step: 30, loss: 0.1948842853307724

step: 40, loss: 0.0502813458442688

step: 50, loss: 0.06158175319433212

step: 60, loss: 0.020530762150883675

step: 70, loss: 0.005139817018061876

```

step: 80, loss: 0.13119915127754211
step: 90, loss: 0.033902402967214584
step: 100, loss: 0.0069306883960962296
step: 110, loss: 0.004737945273518562
step: 120, loss: 0.000703038414940238
step: 130, loss: 0.0009464748436585069
step: 140, loss: 0.0011761946370825171
step: 150, loss: 0.005175583530217409
step: 160, loss: 0.0030315499752759933
step: 170, loss: 0.012415042147040367
step: 180, loss: 0.027164893224835396
step: 190, loss: 0.001949924509972334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.821705426356589, f1=0.8010204081632654,
best_f1=0.8010204081632654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012356045190244913
step: 10, loss: 0.0018366992007941008
step: 20, loss: 0.0034529208205640316
step: 30, loss: 0.06993627548217773
step: 40, loss: 0.03230992332100868
step: 50, loss: 0.054173585027456284
step: 60, loss: 0.03184656426310539
step: 70, loss: 0.0033464718144387007
step: 80, loss: 0.00252557173371315
step: 90, loss: 0.00550858722999692
step: 100, loss: 0.0024847525637596846
step: 110, loss: 0.0019131724257022142
step: 120, loss: 0.15852823853492737
step: 130, loss: 0.004181242547929287
step: 140, loss: 0.2608087658882141
step: 150, loss: 0.005956653971225023
step: 160, loss: 0.011152137070894241
step: 170, loss: 0.048073772341012955
step: 180, loss: 0.028367653489112854
step: 190, loss: 0.0032374500297009945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.8277634961439588, f1=0.8092783505154639,

best_f1=0.8092783505154639

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

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

step: 0, loss: 0.003911919891834259

step: 10, loss: 0.0018049576319754124

step: 20, loss: 0.003915080800652504

step: 30, loss: 0.15631181001663208

step: 40, loss: 0.0076608010567724705

step: 50, loss: 0.02921217679977417

step: 60, loss: 0.0072682262398302555

step: 70, loss: 0.003811767790466547

step: 80, loss: 0.003731934353709221

step: 90, loss: 0.022794010117650032

step: 100, loss: 0.0035848647821694613

step: 110, loss: 0.0029098254162818193

step: 120, loss: 0.0009665843099355698

step: 130, loss: 0.004114825278520584

step: 140, loss: 0.001354596926830709

step: 150, loss: 0.0015034031821414828

step: 160, loss: 0.0050492859445512295

step: 170, loss: 0.0019222090486437082

step: 180, loss: 0.19788411259651184

step: 190, loss: 0.0012953378027305007

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

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

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

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=0.8415584415584416, f1=0.8295165394402036,

best_f1=0.8295165394402036

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

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

step: 0, loss: 0.006025401875376701

step: 10, loss: 0.0008426951826550066

step: 20, loss: 0.010826298035681248

step: 30, loss: 0.011092941276729107

step: 40, loss: 0.0003923581971321255

step: 50, loss: 0.0019420963944867253

step: 60, loss: 0.0008201869786716998

step: 70, loss: 0.00886470451951027

```

step: 80, loss: 0.01435796171426773
step: 90, loss: 0.0012528897495940328
step: 100, loss: 0.0004807878867723048
step: 110, loss: 0.20021219551563263
step: 120, loss: 0.0017473384505137801
step: 130, loss: 0.003582046600058675
step: 140, loss: 0.005965871270745993
step: 150, loss: 0.00906206201761961
step: 160, loss: 0.0010909434640780091
step: 170, loss: 0.0013719202252104878
step: 180, loss: 0.001121007022447884
step: 190, loss: 0.0007612768094986677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8421052631578947, f1=0.8226221079691518,
best_f1=0.8226221079691518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002151301596313715
step: 10, loss: 0.006436305586248636
step: 20, loss: 0.0025498338509351015
step: 30, loss: 0.03087548352777958
step: 40, loss: 0.0010499019408598542
step: 50, loss: 0.002303746063262224
step: 60, loss: 0.04901648685336113
step: 70, loss: 0.03218022733926773
step: 80, loss: 0.001266151200979948
step: 90, loss: 0.0015874761156737804
step: 100, loss: 0.0014505411963909864
step: 110, loss: 0.09625940024852753
step: 120, loss: 0.000590191048104316
step: 130, loss: 0.0005700630135834217
step: 140, loss: 0.002922231564298272
step: 150, loss: 0.00586652010679245
step: 160, loss: 0.00023770047118887305
step: 170, loss: 0.0053891874849796295
step: 180, loss: 0.0007029036642052233
step: 190, loss: 0.014697936363518238
/usr/local/lib/python3.7/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=0.8390501319261214, f1=0.8113695090439276,

```

```

best_f1=0.8226221079691518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005459679290652275
step: 10, loss: 0.0006839674315415323
step: 20, loss: 0.0012820535339415073
step: 30, loss: 0.1155930608510971
step: 40, loss: 0.07475128024816513
step: 50, loss: 0.0004773112013936043
step: 60, loss: 0.002180980984121561
step: 70, loss: 0.0035328688099980354
step: 80, loss: 0.0011441652895882726
step: 90, loss: 0.0009250552393496037
step: 100, loss: 0.0010339075233787298
step: 110, loss: 0.0012410603230819106
step: 120, loss: 0.00295945443212986
step: 130, loss: 0.024836450815200806
step: 140, loss: 0.0027268247213214636
step: 150, loss: 0.000700898643117398
step: 160, loss: 0.0010446637170389295
step: 170, loss: 0.0005157214473001659
step: 180, loss: 0.0003429474600125104
step: 190, loss: 0.18740281462669373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.846153846153846, f1=0.8235294117647058,
best_f1=0.8235294117647058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004008151590824127
step: 10, loss: 0.029992815107107162
step: 20, loss: 0.0006505947094410658
step: 30, loss: 0.000545716262422502
step: 40, loss: 0.005965540651232004
step: 50, loss: 0.00064697558991611
step: 60, loss: 0.0007864407962188125
step: 70, loss: 0.0009579811012372375
step: 80, loss: 0.003235014621168375
step: 90, loss: 0.0001457018224755302
step: 100, loss: 0.00029652283410541713
step: 110, loss: 0.00041120784590020776

```



```

step: 120, loss: 0.0026070415042340755
step: 130, loss: 0.0004708186024799943
step: 140, loss: 0.0003252598980907351
step: 150, loss: 0.0016451890114694834
step: 160, loss: 0.0011362440418452024
step: 170, loss: 0.0002064318978227675
step: 180, loss: 0.0010954096214845777
step: 190, loss: 0.0005577605334110558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8480000000000001, f1=0.8293963254593176,
best_f1=0.8293963254593176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013760443544015288
step: 10, loss: 0.00032715994166210294
step: 20, loss: 0.005443019792437553
step: 30, loss: 0.0002404735132586211
step: 40, loss: 0.0009460955625399947
step: 50, loss: 0.00024849706096574664
step: 60, loss: 0.00017248124640900642
step: 70, loss: 0.00042932614451274276
step: 80, loss: 0.0004065384855493903
step: 90, loss: 0.005622556898742914
step: 100, loss: 0.00046253958134911954
step: 110, loss: 0.0009764070273377001
step: 120, loss: 0.0005714583094231784
step: 130, loss: 0.0005947575555182993
step: 140, loss: 0.0005778680788353086
step: 150, loss: 0.0003600523923523724
step: 160, loss: 0.000591434130910784
step: 170, loss: 0.02532225474715233
step: 180, loss: 0.00039044430013746023
step: 190, loss: 0.0005689202807843685
/usr/local/lib/python3.7/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=0.8446866485013624, f1=0.8310991957104559,
best_f1=0.8293963254593176
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 [00:09, 211.06it/s]
load_f1 = 0.8054054054054054
real_f1 = 0.7957559681697614
733it [00:00, 3171.45it/s]/usr/local/lib/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, 170.05it/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/

```

Downloading: 100% 481/481 [00:00<00:00, 399kB/s]
 Downloading: 100% 899k/899k [00:01<00:00, 829kB/s]
 Downloading: 100% 456k/456k [00:00<00:00, 532kB/s]
 Downloading: 100% 501M/501M [00:07<00:00, 67.7MB/s]
 Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 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: 0.5269624590873718
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.42009612917900085
step: 20, loss: 0.29557228088378906
step: 30, loss: 0.4620771110057831
step: 40, loss: 0.4927154779434204
step: 50, loss: 0.36644992232322693
step: 60, loss: 0.5132538080215454
step: 70, loss: 0.33275306224823
step: 80, loss: 0.2741093039512634
step: 90, loss: 0.19356396794319153
step: 100, loss: 0.1836564540863037
step: 110, loss: 0.49142199754714966
step: 120, loss: 0.34530869126319885
step: 130, loss: 0.31454765796661377
step: 140, loss: 0.3720022737979889
step: 150, loss: 0.3028748333454132
step: 160, loss: 0.4020096957683563
step: 170, loss: 0.3399079144001007
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31110358238220215
step: 10, loss: 0.4551311433315277
step: 20, loss: 0.29566216468811035
step: 30, loss: 0.34739387035369873
step: 40, loss: 0.08220171928405762
step: 50, loss: 0.40506723523139954
step: 60, loss: 0.16184772551059723
step: 70, loss: 0.49142760038375854
step: 80, loss: 0.23907031118869781
step: 90, loss: 0.2413884401321411
step: 100, loss: 0.522047758102417
step: 110, loss: 0.2491445094347
step: 120, loss: 0.2747591435909271
step: 130, loss: 0.5372036099433899
step: 140, loss: 0.5125738978385925
step: 150, loss: 0.435205340385437
step: 160, loss: 0.42577889561653137
step: 170, loss: 0.3897530436515808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.19530876017233126, f1=0.19626615605552894,
best_f1=0.19626615605552894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6014149188995361
step: 10, loss: 0.31677061319351196
step: 20, loss: 0.22398042678833008
step: 30, loss: 0.23857833445072174
step: 40, loss: 0.3596836030483246
step: 50, loss: 0.5870120525360107
step: 60, loss: 0.3292202651500702
step: 70, loss: 0.2536293566226959
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.375016450881958
step: 90, loss: 0.5349245071411133
step: 100, loss: 0.17098616063594818
step: 110, loss: 0.06753361970186234
step: 120, loss: 0.5041235089302063
step: 130, loss: 0.3474062383174896
step: 140, loss: 0.17062747478485107
step: 150, loss: 0.13642197847366333

```

```

step: 160, loss: 0.1501234918832779
step: 170, loss: 0.08633546531200409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.768472906403941, f1=0.7670588235294118,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26789775490760803
step: 10, loss: 0.12845556437969208
step: 20, loss: 0.11846211552619934
step: 30, loss: 0.19572186470031738
step: 40, loss: 0.13208745419979095
step: 50, loss: 0.16848376393318176
step: 60, loss: 0.19681726396083832
step: 70, loss: 0.017750516533851624
step: 80, loss: 0.3286527395248413
step: 90, loss: 0.1931634545326233
step: 100, loss: 0.20673516392707825
step: 110, loss: 0.24758297204971313
step: 120, loss: 0.35300925374031067
step: 130, loss: 0.059058211743831635
step: 140, loss: 0.25321462750434875
step: 150, loss: 0.4305489659309387
step: 160, loss: 0.051096513867378235
step: 170, loss: 0.061798352748155594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8087431693989071, f1=0.8483290488431876,
best_f1=0.8483290488431876
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04718012362718582
step: 10, loss: 0.04103487730026245
step: 20, loss: 0.2117745578289032
step: 30, loss: 0.046165548264980316

```

```

step: 40, loss: 0.160163015127182
step: 50, loss: 0.09208382666110992
step: 60, loss: 0.09048242121934891
step: 70, loss: 0.026628028601408005
step: 80, loss: 0.0010211776243522763
step: 90, loss: 0.043443284928798676
step: 100, loss: 0.019115913659334183
step: 110, loss: 0.09774081408977509
step: 120, loss: 0.08531301468610764
step: 130, loss: 0.005328313447535038
step: 140, loss: 0.10080505907535553
step: 150, loss: 0.1322760283946991
step: 160, loss: 0.008321408182382584
step: 170, loss: 0.011652691289782524
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8395061728395062, f1=0.8899297423887588,
best_f1=0.8899297423887588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08413774520158768
step: 10, loss: 0.06379639357328415
step: 20, loss: 0.09528619050979614
step: 30, loss: 0.0013825793284922838
step: 40, loss: 0.018835699185729027
step: 50, loss: 0.02640124224126339
step: 60, loss: 0.16014251112937927
step: 70, loss: 0.02611652947962284
step: 80, loss: 0.019024834036827087
step: 90, loss: 0.16943946480751038
step: 100, loss: 0.0359993577003479
step: 110, loss: 0.138238325715065
step: 120, loss: 0.042165983468294144
step: 130, loss: 0.00851114746183157
step: 140, loss: 0.02103034034371376
step: 150, loss: 0.013140542432665825
step: 160, loss: 0.06814815104007721
step: 170, loss: 0.008169054053723812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.8689320388349513, f1=0.8967136150234742,
best_f1=0.8967136150234742
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07894957810640335
step: 10, loss: 0.01577320322394371
step: 20, loss: 0.0034674201160669327
step: 30, loss: 0.0018042713636532426
step: 40, loss: 0.0012545263161882758
step: 50, loss: 0.0033349355217069387
step: 60, loss: 0.0045881085097789764
step: 70, loss: 0.004076269455254078
step: 80, loss: 0.07167721539735794
step: 90, loss: 0.13534444570541382
step: 100, loss: 0.007212559226900339
step: 110, loss: 0.05000784620642662
step: 120, loss: 0.03306194022297859
step: 130, loss: 0.0013732612133026123
step: 140, loss: 0.007788517978042364
step: 150, loss: 0.22183166444301605
step: 160, loss: 0.008146794512867928
step: 170, loss: 0.039569295942783356
/usr/local/lib/python3.7/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=0.8656716417910448, f1=0.8794326241134751,
best_f1=0.8967136150234742
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01744999550282955
step: 10, loss: 0.008956272155046463
step: 20, loss: 0.0033302856609225273
step: 30, loss: 0.0010737733682617545
step: 40, loss: 0.008636003360152245
step: 50, loss: 0.020762931555509567
step: 60, loss: 0.03829711303114891
step: 70, loss: 0.0125799048691988
step: 80, loss: 0.0008291505509987473
step: 90, loss: 0.13600485026836395
step: 100, loss: 0.0010288894409313798
step: 110, loss: 0.0327112190425396
step: 120, loss: 0.013873398303985596
step: 130, loss: 0.0015283535467460752
```



```

step: 140, loss: 0.043234046548604965
step: 150, loss: 0.01010804157704115
step: 160, loss: 0.01912744529545307
step: 170, loss: 0.05025447532534599
/usr/local/lib/python3.7/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=0.8615384615384616, f1=0.8963855421686746,
best_f1=0.8967136150234742
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039879210293293
step: 10, loss: 0.0017918971134349704
step: 20, loss: 0.0035523087717592716
step: 30, loss: 0.013829040341079235
step: 40, loss: 0.0030671239364892244
step: 50, loss: 0.0007979085203260183
step: 60, loss: 0.02932099811732769
step: 70, loss: 0.09542224556207657
step: 80, loss: 0.006677941884845495
step: 90, loss: 0.02301989682018757
step: 100, loss: 0.014242606237530708
step: 110, loss: 0.0830879881978035
step: 120, loss: 0.015504532493650913
step: 130, loss: 0.00026920498930849135
step: 140, loss: 0.000619278522208333
step: 150, loss: 0.0667896419763565
step: 160, loss: 0.0011661278549581766
step: 170, loss: 0.003822831204161048
/usr/local/lib/python3.7/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=0.8629441624365481, f1=0.9073634204275535,
best_f1=0.8967136150234742
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007745374459773302
step: 10, loss: 0.008990095928311348
step: 20, loss: 0.04366593062877655
step: 30, loss: 0.07391249388456345
step: 40, loss: 0.0005026143044233322
step: 50, loss: 0.02204935997724533
step: 60, loss: 0.05279068648815155
step: 70, loss: 0.04172050207853317
step: 80, loss: 0.011540045961737633
step: 90, loss: 0.003948595374822617

```

```

step: 100, loss: 0.013317435048520565
step: 110, loss: 0.014315279200673103
step: 120, loss: 0.0003219331556465477
step: 130, loss: 0.005021983291953802
step: 140, loss: 0.0009677385678514838
step: 150, loss: 0.0077218287624418736
step: 160, loss: 0.00012723852705676109
step: 170, loss: 0.021399345248937607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8697788697788698, f1=0.8976744186046511,
best_f1=0.8976744186046511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.050102509558200836
step: 10, loss: 0.0013104331446811557
step: 20, loss: 0.00018394690414424986
step: 30, loss: 0.01099060196429491
step: 40, loss: 0.0017083993880078197
step: 50, loss: 0.012494540773332119
step: 60, loss: 0.018486682325601578
step: 70, loss: 0.00019038902246393263
step: 80, loss: 0.13076435029506683
step: 90, loss: 0.0019174194894731045
step: 100, loss: 0.005064149387180805
step: 110, loss: 0.0006002118461765349
step: 120, loss: 0.0002560388238634914
step: 130, loss: 0.009191679768264294
step: 140, loss: 0.0005975232925266027
step: 150, loss: 0.0001452707074349746
step: 160, loss: 0.011669259518384933
step: 170, loss: 0.010548430494964123
/usr/local/lib/python3.7/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=0.864450127877238, f1=0.8953771289537714,
best_f1=0.8976744186046511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004015340004116297
step: 10, loss: 0.022857461124658585

```

```

step: 20, loss: 0.024193452671170235
step: 30, loss: 0.044746797531843185
step: 40, loss: 0.00033530229120515287
step: 50, loss: 0.00044023734517395496
step: 60, loss: 0.004659837577491999
step: 70, loss: 0.00012143038475187495
step: 80, loss: 0.0001667779724812135
step: 90, loss: 0.023190703243017197
step: 100, loss: 0.00028620538068935275
step: 110, loss: 0.014738512225449085
step: 120, loss: 0.015751458704471588
step: 130, loss: 0.010595639236271381
step: 140, loss: 9.695484186522663e-05
step: 150, loss: 0.0015456498367711902
step: 160, loss: 0.030750222504138947
step: 170, loss: 0.04325816407799721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8765133171912832, f1=0.8930232558139534,
best_f1=0.8930232558139534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016061116475611925
step: 10, loss: 0.00015916689881123602
step: 20, loss: 0.0011229213560000062
step: 30, loss: 0.035850297659635544
step: 40, loss: 0.00034969282569363713
step: 50, loss: 0.006609334610402584
step: 60, loss: 0.0016427168156951666
step: 70, loss: 0.0401369109749794
step: 80, loss: 0.00040292003541253507
step: 90, loss: 0.0012180112535133958
step: 100, loss: 0.031322699040174484
step: 110, loss: 0.00037198886275291443
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 120, loss: 0.05387856811285019
step: 130, loss: 0.00022348783386405557
step: 140, loss: 0.003937280736863613
step: 150, loss: 0.0002314115990884602
step: 160, loss: 0.0002604226174298674
step: 170, loss: 0.0027178179007023573
/usr/local/lib/python3.7/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=0.8675324675324675, f1=0.8967254408060452,
best_f1=0.8930232558139534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008662599138915539
step: 10, loss: 0.0002818634093273431
step: 20, loss: 0.05354475975036621
step: 30, loss: 0.01123605016618967
step: 40, loss: 0.016554560512304306
step: 50, loss: 0.0001556985080242157
step: 60, loss: 0.003540797857567668
step: 70, loss: 0.005099188536405563
step: 80, loss: 0.00016385065100621432
step: 90, loss: 0.00044177763629704714
step: 100, loss: 0.00022439516033045948
step: 110, loss: 0.0240939874202013
step: 120, loss: 0.0015353003982454538
step: 130, loss: 0.014522942714393139
step: 140, loss: 0.0297003835439682
step: 150, loss: 0.025538954883813858
step: 160, loss: 0.00016284068988170475
step: 170, loss: 0.11783705651760101
/usr/local/lib/python3.7/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=0.8630490956072351, f1=0.896551724137931,
best_f1=0.8930232558139534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.896815855521709e-05
step: 10, loss: 0.0001975164923351258
step: 20, loss: 0.0002540151763241738
step: 30, loss: 0.0004881561326328665
step: 40, loss: 0.00016903514915611595
step: 50, loss: 0.0003767718735616654
step: 60, loss: 0.0005025306600145996
step: 70, loss: 0.05706216022372246
step: 80, loss: 0.00023191249056253582
step: 90, loss: 0.0025779150892049074
step: 100, loss: 0.043842609971761703
step: 110, loss: 0.0005546894972212613
step: 120, loss: 0.00011215973790967837
step: 130, loss: 0.0007375045097433031
step: 140, loss: 0.00444828812032938

```

```

step: 150, loss: 0.00010007512173615396
step: 160, loss: 0.00013225294242147356
step: 170, loss: 0.000985535909421742
/usr/local/lib/python3.7/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=0.8693586698337292, f1=0.8904109589041096,
best_f1=0.8930232558139534
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 216.15it/s]
load_f1 = 0.454326923076923
real_f1 = 0.43428571428571433
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 [00:34, 128.38it/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: 0.5985296964645386
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: 0.47480934858322144
step: 20, loss: 0.46314191818237305
step: 30, loss: 0.32133451104164124
step: 40, loss: 0.3730582594871521
step: 50, loss: 0.5298506021499634
step: 60, loss: 0.31459811329841614
step: 70, loss: 0.10515186935663223
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: 0.17199276387691498
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 90, loss: 0.17478372156620026
step: 100, loss: 0.10804172605276108
step: 110, loss: 0.11689542979001999
step: 120, loss: 0.05697052553296089
step: 130, loss: 0.06529844552278519
step: 140, loss: 0.0486677922308445
step: 150, loss: 0.22613295912742615
step: 160, loss: 0.10109642893075943
step: 170, loss: 0.06694784760475159
step: 180, loss: 0.23801253736019135
step: 190, loss: 0.0319223552942276
step: 200, loss: 0.04344755783677101
step: 210, loss: 0.06932993978261948
step: 220, loss: 0.0663265809416771
step: 230, loss: 0.014412306249141693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9601769911504424, f1=0.9594594594594594,

best_f1=0.9594594594594594

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

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

step: 0, loss: 0.01563534326851368

step: 10, loss: 0.18862317502498627

step: 20, loss: 0.015704669058322906

step: 30, loss: 0.06452102214097977

step: 40, loss: 0.1265813708305359

step: 50, loss: 0.010278798639774323

step: 60, loss: 0.012226470746099949

step: 70, loss: 0.009867789223790169

step: 80, loss: 0.010363784618675709

step: 90, loss: 0.07769902795553207

step: 100, loss: 0.0037584647070616484

step: 110, loss: 0.06757950037717819

step: 120, loss: 0.0050148433074355125

step: 130, loss: 0.019780660048127174

step: 140, loss: 0.004368619993329048

step: 150, loss: 0.07744316011667252

step: 160, loss: 0.04342693090438843

step: 170, loss: 0.023936940357089043

step: 180, loss: 0.011327964253723621

step: 190, loss: 0.005113203544169664

step: 200, loss: 0.04420066252350807

step: 210, loss: 0.0304123442620039

step: 220, loss: 0.008969881571829319

step: 230, loss: 0.002978700213134289

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

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

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

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=0.9639344262295082, f1=0.9678135405105438,

best_f1=0.9678135405105438

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

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

step: 0, loss: 0.009595154784619808

step: 10, loss: 0.0026936971116811037

step: 20, loss: 0.02908552810549736

step: 30, loss: 0.0035489555448293686


```

step: 40, loss: 0.009200286120176315
step: 50, loss: 0.007985399104654789
step: 60, loss: 0.009985903277993202
step: 70, loss: 0.009192903526127338
step: 80, loss: 0.13875848054885864
step: 90, loss: 0.01579699106514454
step: 100, loss: 0.012870014645159245
step: 110, loss: 0.05162956565618515
step: 120, loss: 0.04633411020040512
step: 130, loss: 0.009518492966890335
step: 140, loss: 0.0029289452359080315
step: 150, loss: 0.020049424842000008
step: 160, loss: 0.009062394499778748
step: 170, loss: 0.0028268280439078808
step: 180, loss: 0.017736446112394333
step: 190, loss: 0.05512691289186478
step: 200, loss: 0.0074299052357673645
step: 210, loss: 0.010521970689296722
step: 220, loss: 0.11002060770988464
step: 230, loss: 0.016374394297599792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9735682819383259, f1=0.9710467706013363,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006774042267352343
step: 10, loss: 0.0016130972653627396
step: 20, loss: 0.0039045754820108414
step: 30, loss: 0.002529748482629657
step: 40, loss: 0.12159550189971924
step: 50, loss: 0.06141545623540878
step: 60, loss: 0.006170056294649839
step: 70, loss: 0.021034009754657745
step: 80, loss: 0.12953832745552063
step: 90, loss: 0.03370463103055954
step: 100, loss: 0.02214164286851883
step: 110, loss: 0.17916491627693176
step: 120, loss: 0.011788081377744675
step: 130, loss: 0.023286504670977592
step: 140, loss: 0.005460522137582302
step: 150, loss: 0.025469666346907616

```

```

step: 160, loss: 0.004691160749644041
step: 170, loss: 0.0027553376276046038
step: 180, loss: 0.10229631513357162
step: 190, loss: 0.002365718362852931
step: 200, loss: 0.14754198491573334
step: 210, loss: 0.002418826101347804
step: 220, loss: 0.003233996219933033
step: 230, loss: 0.025469906628131866
/usr/local/lib/python3.7/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=0.9732739420935412, f1=0.971815107102593,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002142850309610367
step: 10, loss: 0.007103268057107925
step: 20, loss: 0.16560964286327362
step: 30, loss: 0.003375664120540023
step: 40, loss: 0.003210366703569889
step: 50, loss: 0.004326459486037493
step: 60, loss: 0.022661959752440453
step: 70, loss: 0.004639809485524893
step: 80, loss: 0.017762355506420135
step: 90, loss: 0.24466116726398468
step: 100, loss: 0.00139366346411407
step: 110, loss: 0.005565533880144358
step: 120, loss: 0.019465168938040733
step: 130, loss: 0.0006000573048368096
step: 140, loss: 0.002754490589722991
step: 150, loss: 0.0026825929526239634
step: 160, loss: 0.0038661733269691467
step: 170, loss: 0.08171208202838898
step: 180, loss: 0.00649946928024292
step: 190, loss: 0.1327909529209137
step: 200, loss: 0.015307990834116936
step: 210, loss: 0.013435604982078075
step: 220, loss: 0.006359008140861988
step: 230, loss: 0.021223120391368866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9766407119021134, f1=0.9741282339707535,

```

```

best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013463012874126434
step: 10, loss: 0.002674988703802228
step: 20, loss: 0.0014726452063769102
step: 30, loss: 0.009369200095534325
step: 40, loss: 0.0006168137770146132
step: 50, loss: 0.0016477527096867561
step: 60, loss: 0.0022539435885846615
step: 70, loss: 0.1213204488158226
step: 80, loss: 0.001684657298028469
step: 90, loss: 0.008915971033275127
step: 100, loss: 0.002608510898426175
step: 110, loss: 0.013656968250870705
step: 120, loss: 0.00262182648293674
step: 130, loss: 0.006824670359492302
step: 140, loss: 0.004485061392188072
step: 150, loss: 0.00016967275587376207
step: 160, loss: 0.007938124239444733
step: 170, loss: 0.0013982800301164389
step: 180, loss: 0.02457372471690178
step: 190, loss: 0.017559118568897247
step: 200, loss: 0.02873748540878296
step: 210, loss: 0.02027086168527603
step: 220, loss: 0.04179597645998001
step: 230, loss: 0.0011201922316104174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787709497206705, f1=0.9720044792833147,
best_f1=0.9720044792833147
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031154577154666185
step: 10, loss: 0.004687667824327946
step: 20, loss: 0.0010925219394266605
step: 30, loss: 0.0009528935770504177
step: 40, loss: 0.0007379695889540017
step: 50, loss: 0.013749010860919952
step: 60, loss: 0.0028386006597429514
step: 70, loss: 0.0008371261646971107

```

```

step: 80, loss: 0.0012629777193069458
step: 90, loss: 0.002912412630394101
step: 100, loss: 0.00021676154574379325
step: 110, loss: 0.00045126420445740223
step: 120, loss: 0.0005778245977126062
step: 130, loss: 0.0026207552291452885
step: 140, loss: 0.00022667671146336943
step: 150, loss: 0.05136165767908096
step: 160, loss: 0.00330362725071609
step: 170, loss: 0.004034614190459251
step: 180, loss: 0.001437703613191843
step: 190, loss: 0.002475254237651825
step: 200, loss: 0.00952198076993227
step: 210, loss: 0.0005798250786028802
step: 220, loss: 0.0014308611862361431
step: 230, loss: 0.003564520739018917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.978865406006674, f1=0.9730337078651685,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009105928242206573
step: 10, loss: 0.004630929324775934
step: 20, loss: 0.0005488613387569785
step: 30, loss: 0.0017713672714307904
step: 40, loss: 0.05188167840242386
step: 50, loss: 0.0020567462779581547
step: 60, loss: 0.004533184226602316
step: 70, loss: 0.0016639126697555184
step: 80, loss: 0.17663593590259552
step: 90, loss: 0.0005617028218694031
step: 100, loss: 0.000315046840114519
step: 110, loss: 0.0014320667833089828
step: 120, loss: 0.06615988165140152
step: 130, loss: 0.00090040173381567
step: 140, loss: 0.0002212508989032358
step: 150, loss: 0.1306079477071762
step: 160, loss: 0.0005189395742490888
step: 170, loss: 0.008530831895768642
step: 180, loss: 0.00036258515319786966
step: 190, loss: 0.0013246280141174793

```

```

step: 200, loss: 0.025440484285354614
step: 210, loss: 0.0005395305342972279
step: 220, loss: 0.0004960051155649126
step: 230, loss: 0.000588334514759481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843749999999999, f1=0.9776286353467561,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006421676371246576
step: 10, loss: 0.0006223765667527914
step: 20, loss: 0.0006263076211325824
step: 30, loss: 0.0014201042940840125
step: 40, loss: 0.003283523255959153
step: 50, loss: 0.0002692133712116629
step: 60, loss: 9.668199345469475e-05
step: 70, loss: 0.032676633447408676
step: 80, loss: 0.0006204934907145798
step: 90, loss: 0.05463108420372009
step: 100, loss: 0.009594950824975967
step: 110, loss: 0.002552812220528722
step: 120, loss: 0.023048080503940582
step: 130, loss: 0.003268801374360919
step: 140, loss: 0.00302121601998806
step: 150, loss: 0.0006916624261066318
step: 160, loss: 0.0004747802158817649
step: 170, loss: 0.0013839410385116935
step: 180, loss: 0.001162953325547278
step: 190, loss: 0.00010522198135731742
step: 200, loss: 0.09415888041257858
step: 210, loss: 0.12192559987306595
step: 220, loss: 0.004388167057186365
step: 230, loss: 0.0017593736993148923
/usr/local/lib/python3.7/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=0.9832026875699889, f1=0.9775784753363228,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00047318244469352067
step: 10, loss: 0.00320108188316226
step: 20, loss: 0.00021651617134921253
step: 30, loss: 0.0002092559589073062
step: 40, loss: 0.00021811036276631057
step: 50, loss: 0.00013266864698380232
step: 60, loss: 0.00020623364252969623
step: 70, loss: 0.005022856406867504
step: 80, loss: 8.947874448494986e-05
step: 90, loss: 0.0001080986185115762
step: 100, loss: 0.0007856398588046432
step: 110, loss: 0.0001287877676077187
step: 120, loss: 0.00013204806600697339
step: 130, loss: 0.0002503258001524955
step: 140, loss: 0.0010083048837259412
step: 150, loss: 0.0010112104937434196
step: 160, loss: 6.0367692640284076e-05
step: 170, loss: 0.0008016506908461452
step: 180, loss: 0.0004093602183274925
step: 190, loss: 0.002031659707427025
step: 200, loss: 0.0004003766516689211
step: 210, loss: 0.00025619231746532023
step: 220, loss: 0.0002886317379307002
step: 230, loss: 0.00015885752509348094
/usr/local/lib/python3.7/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=0.9798657718120806,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012571785191539675
step: 10, loss: 0.00016419182065874338
step: 20, loss: 0.0012616260210052133
step: 30, loss: 0.00022366561461240053
step: 40, loss: 0.00012039840657962486
step: 50, loss: 0.0003036763519048691
step: 60, loss: 0.011837628670036793
step: 70, loss: 9.931108070304617e-05
step: 80, loss: 0.0002073944779112935
step: 90, loss: 0.025407426059246063
step: 100, loss: 0.00026142236310988665
step: 110, loss: 0.0003462577878963202
step: 120, loss: 0.000516123662237078
step: 130, loss: 0.00013612843758892268
step: 140, loss: 0.0050653126090765
step: 150, loss: 0.00020806399697903544

```

```

step: 160, loss: 0.012924457900226116
step: 170, loss: 0.00033393377088941634
step: 180, loss: 0.000279881147434935
step: 190, loss: 7.052333967294544e-05
step: 200, loss: 0.006266230251640081
step: 210, loss: 8.940280531533062e-05
step: 220, loss: 0.00019151062588207424
step: 230, loss: 0.0001417869789293036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.984304932735426, f1=0.9808773903262092,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00037348849582485855
step: 10, loss: 0.00010618281521601602
step: 20, loss: 0.015890127047896385
step: 30, loss: 0.034445323050022125
step: 40, loss: 8.999690180644393e-05
step: 50, loss: 0.0001492139999754727
step: 60, loss: 0.00013874712749384344
step: 70, loss: 5.956283712293953e-05
step: 80, loss: 4.13195084547624e-05
step: 90, loss: 0.001701712841168046
step: 100, loss: 9.411212522536516e-05
step: 110, loss: 4.057433034176938e-05
step: 120, loss: 0.0003115905274171382
step: 130, loss: 5.5868491472210735e-05
step: 140, loss: 6.060046507627703e-05
step: 150, loss: 6.656261393800378e-05
step: 160, loss: 5.7653178373584524e-05
step: 170, loss: 9.795393998501822e-05
step: 180, loss: 0.00025998151977546513
step: 190, loss: 0.00015050882939249277
step: 200, loss: 4.5364362449618056e-05
step: 210, loss: 0.0004655691736843437
step: 220, loss: 0.017524704337120056
step: 230, loss: 7.088966958690435e-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=0.9832026875699889, f1=0.9787234042553192,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0001050666396622546
step: 10, loss: 0.00010402878979220986
step: 20, loss: 5.5291438911808655e-05
step: 30, loss: 5.180473090149462e-05
step: 40, loss: 0.0002697290328796953
step: 50, loss: 0.0002462471602484584
step: 60, loss: 9.417304681846872e-05
step: 70, loss: 0.0005942056886851788
step: 80, loss: 0.00012299091031309217
step: 90, loss: 5.345401950762607e-05
step: 100, loss: 0.000826049130409956
step: 110, loss: 0.0001431547134416178
step: 120, loss: 5.783524102298543e-05
step: 130, loss: 5.9094814787385985e-05
step: 140, loss: 0.0009530745446681976
step: 150, loss: 0.00024693162413313985
step: 160, loss: 0.013552446849644184
step: 170, loss: 6.395179661922157e-05
step: 180, loss: 0.037583477795124054
step: 190, loss: 7.369283412117511e-05
step: 200, loss: 0.009841189719736576
step: 210, loss: 0.00027140011661686003
step: 220, loss: 0.0001111346937250346
step: 230, loss: 8.258514571934938e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9854423292273236, f1=0.9808342728297633,
best_f1=0.9808342728297633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002146780170733109
step: 10, loss: 5.719167165807448e-05
step: 20, loss: 7.71923951106146e-05
step: 30, loss: 8.366911788471043e-05
step: 40, loss: 5.113822771818377e-05
step: 50, loss: 0.0006150571862235665
step: 60, loss: 5.804406828247011e-05
step: 70, loss: 7.361212919931859e-05
step: 80, loss: 0.0008244143682532012
step: 90, loss: 0.0001387272059218958
step: 100, loss: 0.00014796435425523669
step: 110, loss: 6.895129627082497e-05

```



```

step: 120, loss: 3.599283081712201e-05
step: 130, loss: 0.000190396182006225
step: 140, loss: 0.00021645576634909958
step: 150, loss: 8.879349479684606e-05
step: 160, loss: 6.950111128389835e-05
step: 170, loss: 0.00012788211461156607
step: 180, loss: 5.332075306796469e-05
step: 190, loss: 4.249233825248666e-05
step: 200, loss: 0.0013620815007016063
step: 210, loss: 4.795753557118587e-05
step: 220, loss: 0.0001234694354934618
step: 230, loss: 0.006599004380404949
/usr/local/lib/python3.7/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=0.9832402234636871, f1=0.9797752808988766,
best_f1=0.9808342728297633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024577539414167404
step: 10, loss: 4.365142376627773e-05
step: 20, loss: 0.00019455014262348413
step: 30, loss: 6.52778908261098e-05
step: 40, loss: 3.9028687751851976e-05
step: 50, loss: 4.416133015183732e-05
step: 60, loss: 0.030808907002210617
step: 70, loss: 0.001531400135718286
step: 80, loss: 9.159908222500235e-05
step: 90, loss: 5.614152178168297e-05
step: 100, loss: 4.399407407618128e-05
step: 110, loss: 8.66408008732833e-05
step: 120, loss: 0.03353222459554672
step: 130, loss: 0.00012233089364599437
step: 140, loss: 0.010040992870926857
step: 150, loss: 0.00013549940194934607
step: 160, loss: 0.0003212092851754278
step: 170, loss: 3.8696976844221354e-05
step: 180, loss: 5.977703403914347e-05
step: 190, loss: 0.00020989132463000715
step: 200, loss: 8.146007894538343e-05
step: 210, loss: 0.01797565445303917
step: 220, loss: 4.937983976560645e-05
step: 230, loss: 3.8298676372505724e-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=0.9832402234636871, f1=0.9808342728297633,

```

```

best_f1=0.9808342728297633
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:13, 183.83it/s]
load_f1 = 0.9855072463768116
real_f1 = 0.9843749999999999
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 [00:25, 169.84it/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: 0.7002458572387695
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4296877980232239
step: 20, loss: 0.3745748698711395
step: 30, loss: 0.3507492244243622
step: 40, loss: 0.35930952429771423
step: 50, loss: 0.5788509249687195
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.24378320574760437
step: 70, loss: 0.2640804648399353
step: 80, loss: 0.1134454607963562
step: 90, loss: 0.23873311281204224
step: 100, loss: 0.2065601348876953
step: 110, loss: 0.12639984488487244
step: 120, loss: 0.16591928899288177
step: 130, loss: 0.19465214014053345
step: 140, loss: 0.3335929214954376
step: 150, loss: 0.1347513347864151
step: 160, loss: 0.17865338921546936
step: 170, loss: 0.079416424036026
step: 180, loss: 0.045789144933223724
step: 190, loss: 0.08783650398254395
step: 200, loss: 0.03693623095750809
step: 210, loss: 0.035847779363393784
step: 220, loss: 0.009577409364283085
step: 230, loss: 0.09974724799394608
step: 240, loss: 0.011688456870615482
step: 250, loss: 0.20579059422016144
step: 260, loss: 0.16999763250350952
step: 270, loss: 0.4382694959640503
step: 280, loss: 0.025372611358761787
step: 290, loss: 0.08413532376289368

```

```

step: 300, loss: 0.11513668298721313
step: 310, loss: 0.11266113072633743
step: 320, loss: 0.20291204750537872
step: 330, loss: 0.06751095503568649
step: 340, loss: 0.30556946992874146
step: 350, loss: 0.1062081903219223
step: 360, loss: 0.08835334330797195
step: 370, loss: 0.01532131526619196
step: 380, loss: 0.13596917688846588
step: 390, loss: 0.010900688357651234
step: 400, loss: 0.056525733321905136
step: 410, loss: 0.28162097930908203
step: 420, loss: 0.020960889756679535
step: 430, loss: 0.030355751514434814
step: 440, loss: 0.004686322528868914
step: 450, loss: 0.0463327057659626
step: 460, loss: 0.04200833663344383
step: 470, loss: 0.08933047950267792
step: 480, loss: 0.1883012354373932
step: 490, loss: 0.14132550358772278
step: 500, loss: 0.03924933820962906
step: 510, loss: 0.032781220972537994
step: 520, loss: 0.021738866344094276
step: 530, loss: 0.004355949815362692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9330819981149858, f1=0.937207122774133,
best_f1=0.937207122774133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08573666960000992
step: 10, loss: 0.06840604543685913
step: 20, loss: 0.02332344837486744
step: 30, loss: 0.04650597274303436
step: 40, loss: 0.08804238587617874
step: 50, loss: 0.06217392906546593
step: 60, loss: 0.067331962287426
step: 70, loss: 0.013451937586069107
step: 80, loss: 0.032791078090667725
step: 90, loss: 0.024235408753156662
step: 100, loss: 0.12044985592365265
step: 110, loss: 0.03195923939347267

```

```
step: 120, loss: 0.017643246799707413
step: 130, loss: 0.01409414317458868
step: 140, loss: 0.11847928911447525
step: 150, loss: 0.028346911072731018
step: 160, loss: 0.0353870764374733
step: 170, loss: 0.06548872590065002
step: 180, loss: 0.08418828994035721
step: 190, loss: 0.017044810578227043
step: 200, loss: 0.0670410618185997
step: 210, loss: 0.02070363610982895
step: 220, loss: 0.001526020118035376
step: 230, loss: 0.08870882540941238
step: 240, loss: 0.04757476970553398
step: 250, loss: 0.004266396630555391
step: 260, loss: 0.20044878125190735
step: 270, loss: 0.010728185065090656
step: 280, loss: 0.03936106339097023
step: 290, loss: 0.05309903621673584
step: 300, loss: 0.056492675095796585
step: 310, loss: 0.021621450781822205
step: 320, loss: 0.017465243116021156
step: 330, loss: 0.04790763556957245
step: 340, loss: 0.11667274683713913
step: 350, loss: 0.0025976679753512144
step: 360, loss: 0.09603127092123032
step: 370, loss: 0.002664949744939804
step: 380, loss: 0.06534889340400696
step: 390, loss: 0.009524619206786156
step: 400, loss: 0.08463460206985474
step: 410, loss: 0.051431622356176376
step: 420, loss: 0.08393795043230057
step: 430, loss: 0.25309404730796814
step: 440, loss: 0.007569790352135897
step: 450, loss: 0.008423130959272385
step: 460, loss: 0.0474131740629673
step: 470, loss: 0.009578024968504906
step: 480, loss: 0.0072915200144052505
step: 490, loss: 0.023481065407395363
step: 500, loss: 0.011172506026923656
step: 510, loss: 0.06724458932876587
step: 520, loss: 0.2982431650161743
step: 530, loss: 0.052707064896821976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9517625231910947, f1=0.9515828677839852,
best_f1=0.9515828677839852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1417747437953949
step: 10, loss: 0.050325166434049606
step: 20, loss: 0.008190945722162724
step: 30, loss: 0.11180978268384933
step: 40, loss: 0.06055290624499321
step: 50, loss: 0.008407960645854473
step: 60, loss: 0.017159773036837578
step: 70, loss: 0.013784460723400116
step: 80, loss: 0.1469993144273758
step: 90, loss: 0.004404775332659483
step: 100, loss: 0.02367861568927765
step: 110, loss: 0.006088870111852884
step: 120, loss: 0.2933160662651062
step: 130, loss: 0.05984184145927429
step: 140, loss: 0.02425006777048111
step: 150, loss: 0.0416024848818779
step: 160, loss: 0.09942873567342758
step: 170, loss: 0.0015803944552317262
step: 180, loss: 0.0037136049941182137
step: 190, loss: 0.0029369560070335865
step: 200, loss: 0.008096577599644661
step: 210, loss: 0.04300391301512718
step: 220, loss: 0.10642237961292267
step: 230, loss: 0.024918420240283012
step: 240, loss: 0.026928046718239784
step: 250, loss: 0.017232896760106087
step: 260, loss: 0.10599192976951599
step: 270, loss: 0.0022410559467971325
step: 280, loss: 0.0024151757825165987
step: 290, loss: 0.00784221850335598
step: 300, loss: 0.15274304151535034
step: 310, loss: 0.05294668674468994
step: 320, loss: 0.016438014805316925
step: 330, loss: 0.07211791723966599
step: 340, loss: 0.0072308010421693325
step: 350, loss: 0.053566060960292816
step: 360, loss: 0.024482375010848045
step: 370, loss: 0.015494752675294876
step: 380, loss: 0.0037717653904110193
step: 390, loss: 0.008654329925775528
step: 400, loss: 0.09219314157962799
step: 410, loss: 0.04201538488268852
```

```

step: 420, loss: 0.010818001814186573
step: 430, loss: 0.013327211141586304
step: 440, loss: 0.13680000603199005
step: 450, loss: 0.008718234486877918
step: 460, loss: 0.02932986430823803
step: 470, loss: 0.019783923402428627
step: 480, loss: 0.15600763261318207
step: 490, loss: 0.034010060131549835
step: 500, loss: 0.10617798566818237
step: 510, loss: 0.09642846882343292
step: 520, loss: 0.03573213145136833
step: 530, loss: 0.0029379904735833406
/usr/local/lib/python3.7/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=0.9475138121546962, f1=0.9507434944237918,
best_f1=0.9515828677839852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043199196457862854
step: 10, loss: 0.008930772542953491
step: 20, loss: 0.01661134883761406
step: 30, loss: 0.13807062804698944
step: 40, loss: 0.006616718601435423
step: 50, loss: 0.09224992990493774
step: 60, loss: 0.0012475951807573438
step: 70, loss: 0.034692175686359406
step: 80, loss: 0.18091611564159393
step: 90, loss: 0.1027030348777771
step: 100, loss: 0.007356897462159395
step: 110, loss: 0.05285205692052841
step: 120, loss: 0.0026589955668896437
step: 130, loss: 0.02983604185283184
step: 140, loss: 0.019618578255176544
step: 150, loss: 0.0012168239336460829
step: 160, loss: 0.004315580707043409
step: 170, loss: 0.011110451072454453
step: 180, loss: 0.02071481943130493
step: 190, loss: 0.08571772277355194
step: 200, loss: 0.0718310996890068
step: 210, loss: 0.006212986074388027
step: 220, loss: 0.0018590748077258468
step: 230, loss: 0.0024711426813155413
step: 240, loss: 0.0627722218632698
step: 250, loss: 0.09862492978572845
step: 260, loss: 0.03284228593111038
step: 270, loss: 0.06415899097919464

```



```

step: 280, loss: 0.008622257970273495
step: 290, loss: 0.04390161857008934
step: 300, loss: 0.0007323396857827902
step: 310, loss: 0.016618436202406883
step: 320, loss: 0.03831962123513222
step: 330, loss: 0.025861894711852074
step: 340, loss: 0.015013596974313259
step: 350, loss: 0.103089839220047
step: 360, loss: 0.052028268575668335
step: 370, loss: 0.0008554071537218988
step: 380, loss: 0.001708175870589912
step: 390, loss: 0.0007032349240034819
step: 400, loss: 0.0025987643748521805
step: 410, loss: 0.003577422583475709
step: 420, loss: 0.02000472880899906
step: 430, loss: 0.003712644334882498
step: 440, loss: 0.0381682887673378
step: 450, loss: 0.006434087175875902
step: 460, loss: 0.0036472314968705177
step: 470, loss: 0.05778111517429352
step: 480, loss: 0.08965358138084412
step: 490, loss: 0.11300872266292572
step: 500, loss: 0.03534994274377823
step: 510, loss: 0.03186589479446411
step: 520, loss: 0.016901129856705666
step: 530, loss: 0.19089730083942413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.952919020715631, f1=0.9430740037950663,
best_f1=0.9430740037950663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001972238766029477
step: 10, loss: 0.06856199353933334
step: 20, loss: 0.0028025482315570116
step: 30, loss: 0.016396882012486458
step: 40, loss: 0.00019166740821674466
step: 50, loss: 0.012872700579464436
step: 60, loss: 0.05413976311683655
step: 70, loss: 0.0016676043160259724
step: 80, loss: 0.006574344355612993
step: 90, loss: 0.0023146404419094324

```

```
step: 100, loss: 0.06362899392843246
step: 110, loss: 0.005809334106743336
step: 120, loss: 0.14020057022571564
step: 130, loss: 0.00703634275123477
step: 140, loss: 0.0005607508355751634
step: 150, loss: 0.010794774629175663
step: 160, loss: 0.0018913396634161472
step: 170, loss: 0.09603048861026764
step: 180, loss: 0.04995207488536835
step: 190, loss: 0.030371276661753654
step: 200, loss: 0.0034254053607583046
step: 210, loss: 0.005943878088146448
step: 220, loss: 0.005397643428295851
step: 230, loss: 0.0026182271540164948
step: 240, loss: 0.035353273153305054
step: 250, loss: 0.11827881634235382
step: 260, loss: 0.0003689041768666357
step: 270, loss: 0.03344475477933884
step: 280, loss: 0.04009925574064255
step: 290, loss: 0.002849231706932187
step: 300, loss: 0.14730596542358398
step: 310, loss: 0.06691087037324905
step: 320, loss: 0.046451255679130554
step: 330, loss: 0.003032909706234932
step: 340, loss: 0.009082113392651081
step: 350, loss: 0.0006233786116354167
step: 360, loss: 0.00024004880106076598
step: 370, loss: 0.0018437764374539256
step: 380, loss: 0.00010498673509573564
step: 390, loss: 0.04682526737451553
step: 400, loss: 0.005776938982307911
step: 410, loss: 0.06581049412488937
step: 420, loss: 0.22521081566810608
step: 430, loss: 0.002639236394315958
step: 440, loss: 0.014358392916619778
step: 450, loss: 0.009689169004559517
step: 460, loss: 0.0020338178146630526
step: 470, loss: 0.13652685284614563
step: 480, loss: 0.012463415041565895
step: 490, loss: 0.0024788070004433393
step: 500, loss: 0.006785337347537279
step: 510, loss: 0.0011635483242571354
step: 520, loss: 0.06028824672102928
step: 530, loss: 0.028279460966587067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9586623316302832, f1=0.9505135387488329,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031921397894620895
step: 10, loss: 0.0009477121056988835
step: 20, loss: 0.000920152582693845
step: 30, loss: 0.0006413086666725576
step: 40, loss: 0.005047118756920099
step: 50, loss: 0.00021041592117398977
step: 60, loss: 0.0007720732246525586
step: 70, loss: 0.0010100847575813532
step: 80, loss: 0.00174482143484056
step: 90, loss: 0.01767541468143463
step: 100, loss: 0.004468927159905434
step: 110, loss: 0.019404390826821327
step: 120, loss: 0.012348132207989693
step: 130, loss: 0.004297660198062658
step: 140, loss: 0.0011023711413145065
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.00028526742244139314
step: 160, loss: 0.07974362373352051
step: 170, loss: 0.11660873889923096
step: 180, loss: 0.001471838098950684
step: 190, loss: 0.08096162229776382
step: 200, loss: 0.005965999327600002
step: 210, loss: 0.009547526016831398
step: 220, loss: 0.004332810174673796
step: 230, loss: 0.003579886630177498
step: 240, loss: 0.00025060243206098676
step: 250, loss: 0.06497910618782043
step: 260, loss: 0.0003740307583939284
step: 270, loss: 0.0034342254512012005
step: 280, loss: 0.0030061467550694942
step: 290, loss: 0.0018282230012118816
step: 300, loss: 0.03159680962562561
step: 310, loss: 0.03376837819814682
step: 320, loss: 9.936989226844162e-05
step: 330, loss: 0.0007758995052427053
step: 340, loss: 0.0004515174659900367
step: 350, loss: 0.0008775254827924073
step: 360, loss: 0.11078324913978577
step: 370, loss: 0.00783323310315609
step: 380, loss: 0.0006441588629968464
```

```
step: 390, loss: 0.008870676159858704
step: 400, loss: 0.0028658772353082895
step: 410, loss: 0.0026534374337643385
step: 420, loss: 0.000860006723087281
step: 430, loss: 0.0017400756478309631
step: 440, loss: 0.0021652504801750183
step: 450, loss: 0.23686987161636353
step: 460, loss: 0.004035635851323605
step: 470, loss: 0.005308314226567745
step: 480, loss: 0.021773992106318474
step: 490, loss: 0.005412269849330187
step: 500, loss: 0.001487748115323484
step: 510, loss: 0.06457579135894775
step: 520, loss: 0.0012047933414578438
step: 530, loss: 0.005586845334619284
/usr/local/lib/python3.7/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=0.9565217391304348, f1=0.9475150952159777,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016581424279138446
step: 10, loss: 0.00031355448300018907
step: 20, loss: 0.008876548148691654
step: 30, loss: 0.016045793890953064
step: 40, loss: 0.007632856722921133
step: 50, loss: 0.006378604099154472
step: 60, loss: 0.008668973110616207
step: 70, loss: 0.0011093586217612028
step: 80, loss: 0.00029055416234768927
step: 90, loss: 0.00013264910376165062
step: 100, loss: 0.0007354654371738434
step: 110, loss: 0.0011633398244157434
step: 120, loss: 0.0006961559411138296
step: 130, loss: 9.273529576603323e-05
step: 140, loss: 0.00047599742538295686
step: 150, loss: 0.0020688704680651426
step: 160, loss: 0.00010421469050925225
step: 170, loss: 0.046938907355070114
step: 180, loss: 0.0006066677742637694
step: 190, loss: 0.0013204620918259025
step: 200, loss: 0.00042207184014841914
step: 210, loss: 0.00011079633259214461
step: 220, loss: 0.026224453002214432
step: 230, loss: 0.00012125306238885969
step: 240, loss: 0.000754091830458492
```

```

step: 250, loss: 0.0019932924769818783
step: 260, loss: 0.0039020597469061613
step: 270, loss: 0.0028797758277505636
step: 280, loss: 0.00694679981097579
step: 290, loss: 0.005075589753687382
step: 300, loss: 0.026069972664117813
step: 310, loss: 0.0009166540112346411
step: 320, loss: 0.00044320645974949
step: 330, loss: 0.00021919973369222134
step: 340, loss: 0.031053422018885612
step: 350, loss: 0.00036086500040255487
step: 360, loss: 0.003456017468124628
step: 370, loss: 0.0021647256799042225
step: 380, loss: 0.006336160469800234
step: 390, loss: 0.0005464269197545946
step: 400, loss: 0.023662487044930458
step: 410, loss: 0.00016613135812804103
step: 420, loss: 0.11885806173086166
step: 430, loss: 0.011607635766267776
step: 440, loss: 0.0040095155127346516
step: 450, loss: 0.0011215365957468748
step: 460, loss: 0.014028836973011494
step: 470, loss: 0.09792967140674591
step: 480, loss: 0.012278336100280285
step: 490, loss: 0.002503792755305767
step: 500, loss: 0.0025807323399931192
step: 510, loss: 0.00023945505381561816
step: 520, loss: 0.0002105186867993325
step: 530, loss: 0.0003280005184933543
/usr/local/lib/python3.7/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=0.9573679332715477, f1=0.9490740740740741,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007314437534660101
step: 10, loss: 0.0021900704596191645
step: 20, loss: 0.003224405227228999
step: 30, loss: 0.0001758302969392389
step: 40, loss: 7.909526175353676e-05
step: 50, loss: 0.0003735884674824774
step: 60, loss: 0.0002415073395241052
step: 70, loss: 0.0005653350963257253
step: 80, loss: 0.02612614445388317
step: 90, loss: 0.0005478406674228609
step: 100, loss: 0.005466769449412823

```

```
step: 110, loss: 0.0002591333177406341
step: 120, loss: 0.0009093886474147439
step: 130, loss: 0.0011009329464286566
step: 140, loss: 0.000277963001281023
step: 150, loss: 0.0008243609918281436
step: 160, loss: 8.139111014315858e-05
step: 170, loss: 0.03129762038588524
step: 180, loss: 0.0001887403050204739
step: 190, loss: 0.00013639690587297082
step: 200, loss: 0.008366030640900135
step: 210, loss: 0.05508317053318024
step: 220, loss: 6.544103234773502e-05
step: 230, loss: 0.03734569624066353
step: 240, loss: 0.02620660699903965
step: 250, loss: 0.0025452764239162207
step: 260, loss: 7.899126649135724e-05
step: 270, loss: 0.0010517026530578732
step: 280, loss: 6.435933755710721e-05
step: 290, loss: 0.0031655237544327974
step: 300, loss: 4.6587996621383354e-05
step: 310, loss: 0.00730480533093214
step: 320, loss: 0.0008030512253753841
step: 330, loss: 7.284696766873822e-05
step: 340, loss: 0.0025448461528867483
step: 350, loss: 3.810876660281792e-05
step: 360, loss: 0.01961403898894787
step: 370, loss: 0.04216599464416504
step: 380, loss: 0.0001958659995580092
step: 390, loss: 0.027579685673117638
step: 400, loss: 0.007807510439306498
step: 410, loss: 0.002379618352279067
step: 420, loss: 0.0011907877633348107
step: 430, loss: 0.0008742212085053325
step: 440, loss: 0.0020569018088281155
step: 450, loss: 0.0010719024576246738
step: 460, loss: 0.0024298513308167458
step: 470, loss: 0.06381392478942871
step: 480, loss: 0.003025804413482547
step: 490, loss: 0.003068941179662943
step: 500, loss: 0.0001482063962612301
step: 510, loss: 0.006557428278028965
step: 520, loss: 0.00017767734243534505
step: 530, loss: 0.0006934028351679444
/usr/local/lib/python3.7/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=0.9546940681924334, f1=0.9514472455648927,
best_f1=0.9505135387488329
```

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

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

step: 0, loss: 0.00012589212565217167
step: 10, loss: 0.00030529891955666244
step: 20, loss: 0.00012723880354315042
step: 30, loss: 0.11616145819425583
step: 40, loss: 0.00037867174251005054
step: 50, loss: 0.002018648898229003
step: 60, loss: 0.0008488207822665572
step: 70, loss: 0.0016973692690953612
step: 80, loss: 0.002917597070336342
step: 90, loss: 0.10960374027490616
step: 100, loss: 0.0014946075389161706
step: 110, loss: 0.008939893916249275
step: 120, loss: 0.000352987990481779
step: 130, loss: 0.0003304198617115617
step: 140, loss: 0.00029391003772616386
step: 150, loss: 0.01598784513771534
step: 160, loss: 5.384163887356408e-05
step: 170, loss: 0.011615785770118237
step: 180, loss: 0.00046334281796589494
step: 190, loss: 5.433671685750596e-05
step: 200, loss: 4.111898670089431e-05
step: 210, loss: 0.002745090750977397
step: 220, loss: 0.003320110496133566
step: 230, loss: 0.0020598170813173056
step: 240, loss: 8.049558528000489e-05
step: 250, loss: 0.00016043504001572728
step: 260, loss: 0.00024577692965976894
step: 270, loss: 0.00026835172320716083
step: 280, loss: 0.004587004892528057
step: 290, loss: 0.008436751551926136
step: 300, loss: 0.00047515268670395017
step: 310, loss: 0.014564030803740025
step: 320, loss: 0.0021826138254255056
step: 330, loss: 0.00010113314783666283
step: 340, loss: 0.04700641706585884
step: 350, loss: 0.00839791540056467
step: 360, loss: 3.195026874891482e-05
step: 370, loss: 3.9329312130576e-05
step: 380, loss: 0.001510533387772739
step: 390, loss: 0.0004395748255774379
step: 400, loss: 6.05440181971062e-05
step: 410, loss: 0.00044685753528028727
step: 420, loss: 2.723042416619137e-05
step: 430, loss: 0.012425730936229229
step: 440, loss: 0.00021019972336944193

```

step: 450, loss: 0.005884942132979631
step: 460, loss: 0.003026658436283469
step: 470, loss: 0.00011472940241219476
step: 480, loss: 0.06246693432331085
step: 490, loss: 0.008567866869270802
step: 500, loss: 0.00043363773147575557
step: 510, loss: 0.0014853544998914003
step: 520, loss: 0.019893109798431396
step: 530, loss: 0.006299149710685015
/usr/local/lib/python3.7/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=0.9543958627174425, f1=0.947219604147031,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021734939946327358
step: 10, loss: 0.0009844829328358173
step: 20, loss: 8.319469634443521e-05
step: 30, loss: 0.0007600535172969103
step: 40, loss: 0.00016554450849071145
step: 50, loss: 0.00018299638759344816
step: 60, loss: 0.0044248346239328384
step: 70, loss: 0.0011740473564714193
step: 80, loss: 0.003017725422978401
step: 90, loss: 0.0006724639097228646
step: 100, loss: 0.00022005653590895236
step: 110, loss: 0.019674891605973244
step: 120, loss: 6.707636930514127e-05
step: 130, loss: 0.0010170588502660394
step: 140, loss: 0.001073441351763904
step: 150, loss: 0.0025367699563503265
step: 160, loss: 0.039040207862854004
step: 170, loss: 0.023354314267635345
step: 180, loss: 0.04747159406542778
step: 190, loss: 0.00040256560896523297
step: 200, loss: 0.00779802817851305
step: 210, loss: 0.011429697275161743
step: 220, loss: 0.0002602752356324345
step: 230, loss: 0.0006244678515940905
step: 240, loss: 8.723077917238697e-05
step: 250, loss: 0.010757491923868656
step: 260, loss: 8.70087169460021e-05
step: 270, loss: 0.0009706052951514721
step: 280, loss: 0.0036871552001684904
step: 290, loss: 0.0007478068000636995
step: 300, loss: 0.0004865251830779016

```



```

step: 310, loss: 0.03543583303689957
step: 320, loss: 0.004492016509175301
step: 330, loss: 0.0008044152054935694
step: 340, loss: 4.857536623603664e-05
step: 350, loss: 0.00037705583963543177
step: 360, loss: 6.647453119512647e-05
step: 370, loss: 3.640945578808896e-05
step: 380, loss: 0.0005731711280532181
step: 390, loss: 7.812760304659605e-05
step: 400, loss: 0.00011845735571114346
step: 410, loss: 0.005616620648652315
step: 420, loss: 3.0073388188611716e-05
step: 430, loss: 0.0002650547830853611
step: 440, loss: 0.0001118276923079975
step: 450, loss: 0.01445856038480997
step: 460, loss: 0.00022923137294128537
step: 470, loss: 0.0015518979635089636
step: 480, loss: 0.00020692261750809848
step: 490, loss: 0.0071977851912379265
step: 500, loss: 8.30718781799078e-05
step: 510, loss: 2.961816790048033e-05
step: 520, loss: 0.002597643993794918
step: 530, loss: 0.0006524226628243923
/usr/local/lib/python3.7/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=0.9574964969640355, f1=0.9468283582089553,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7186226361664012e-05
step: 10, loss: 0.0003285847487859428
step: 20, loss: 0.00037213097675703466
step: 30, loss: 0.00010158589429920539
step: 40, loss: 0.0001655583910178393
step: 50, loss: 0.00024625632795505226
step: 60, loss: 0.000509467558003962
step: 70, loss: 3.5157678212272e-05
step: 80, loss: 0.001610096893273294
step: 90, loss: 0.0026593084912747145
step: 100, loss: 0.0005034137284383178
step: 110, loss: 0.0008671811665408313
step: 120, loss: 0.00032138111419044435
step: 130, loss: 5.058239185018465e-05
step: 140, loss: 0.00015711237210780382
step: 150, loss: 0.0012509243097156286
step: 160, loss: 0.00021886607282795012

```

```

step: 170, loss: 2.9721742976107635e-05
step: 180, loss: 3.7940353649901226e-05
step: 190, loss: 4.873682701145299e-05
step: 200, loss: 2.283901812916156e-05
step: 210, loss: 0.00011357789480825886
step: 220, loss: 0.0020651693921536207
step: 230, loss: 2.09614445338957e-05
step: 240, loss: 0.0037719032261520624
step: 250, loss: 0.0013107319828122854
step: 260, loss: 0.0019289038609713316
step: 270, loss: 0.0010263859294354916
step: 280, loss: 0.028709745034575462
step: 290, loss: 0.001460138475522399
step: 300, loss: 6.075463898014277e-05
step: 310, loss: 0.03310171514749527
step: 320, loss: 0.00011057453957619146
step: 330, loss: 2.6336823793826625e-05
step: 340, loss: 0.002809268655255437
step: 350, loss: 0.00014160822320263833
step: 360, loss: 0.0003583406505640596
step: 370, loss: 0.0022767914924770594
step: 380, loss: 0.0003753795172087848
step: 390, loss: 0.0012654394377022982
step: 400, loss: 6.966434011701494e-05
step: 410, loss: 0.0007105889962986112
step: 420, loss: 0.00826089084148407
step: 430, loss: 0.005819762125611305
step: 440, loss: 0.000411299493862316
step: 450, loss: 0.0010760523146018386
step: 460, loss: 7.671841012779623e-05
step: 470, loss: 0.0002944949665106833
step: 480, loss: 0.0002017343940678984
step: 490, loss: 3.9665384974796325e-05
step: 500, loss: 0.00017659645527601242
step: 510, loss: 0.00012812796921934932
step: 520, loss: 6.247842247830704e-05
step: 530, loss: 0.0223955437541008
/usr/local/lib/python3.7/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=0.9561975768872321, f1=0.9477611940298507,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000852595258038491
step: 10, loss: 3.73344992112834e-05
step: 20, loss: 0.0006329738534986973

```

step: 30, loss: 0.00011752398131648079
step: 40, loss: 0.00012785400031134486
step: 50, loss: 0.0010763216996565461
step: 60, loss: 3.754096542252228e-05
step: 70, loss: 0.0005994639359414577
step: 80, loss: 0.0003597138565964997
step: 90, loss: 6.492229294963181e-05
step: 100, loss: 0.08190605044364929
step: 110, loss: 0.00012032641825499013
step: 120, loss: 0.0003056470595765859
step: 130, loss: 0.0002264681097585708
step: 140, loss: 1.7899639715324156e-05
step: 150, loss: 0.00012069566582795233
step: 160, loss: 0.00024238215701188892
step: 170, loss: 0.00012241938384249806
step: 180, loss: 3.032047243323177e-05
step: 190, loss: 8.371955482289195e-05
step: 200, loss: 0.0012826278107240796
step: 210, loss: 0.0005825177649967372
step: 220, loss: 0.0004284657188691199
step: 230, loss: 1.4744501640961971e-05
step: 240, loss: 0.0015341059770435095
step: 250, loss: 3.023674616997596e-05
step: 260, loss: 1.3548674360208679e-05
step: 270, loss: 0.0005108496407046914
step: 280, loss: 0.00016947482072282583
step: 290, loss: 6.192610453581437e-05
step: 300, loss: 0.0003712469479069114
step: 310, loss: 1.8424880181555636e-05
step: 320, loss: 0.00018715871556196362
step: 330, loss: 0.0006642410880886018
step: 340, loss: 0.00014810329594183713
step: 350, loss: 7.553792966064066e-05
step: 360, loss: 0.0046459692530334
step: 370, loss: 0.00021591542463283986
step: 380, loss: 0.00011980918498011306
step: 390, loss: 0.0017966199666261673
step: 400, loss: 1.3485362615028862e-05
step: 410, loss: 0.008613824844360352
step: 420, loss: 6.226691766642034e-05
step: 430, loss: 4.638710379367694e-05
step: 440, loss: 0.00023908758885227144
step: 450, loss: 5.970041092950851e-05
step: 460, loss: 0.0001797565637389198
step: 470, loss: 0.0013885919470340014
step: 480, loss: 0.019152769818902016
step: 490, loss: 0.0034635053016245365
step: 500, loss: 1.477036585129099e-05

```

step: 510, loss: 0.007369156461209059
step: 520, loss: 0.0011112323263660073
step: 530, loss: 0.00045905375736765563
/usr/local/lib/python3.7/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=0.9570041608876559, f1=0.9485873089393237,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.3105375728628132e-05
step: 10, loss: 4.466096652322449e-05
step: 20, loss: 1.6350018995581195e-05
step: 30, loss: 6.0992570070084184e-05
step: 40, loss: 0.0001192333729704842
step: 50, loss: 0.002182908356189728
step: 60, loss: 8.271427941508591e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.0002554628299549222
step: 80, loss: 1.4774293958907947e-05
step: 90, loss: 0.0009334433707408607
step: 100, loss: 1.717312989057973e-05
step: 110, loss: 1.0896364983636886e-05
step: 120, loss: 1.0620689863571897e-05
step: 130, loss: 9.149240213446319e-06
step: 140, loss: 8.561469439882785e-05
step: 150, loss: 0.0017953513888642192
step: 160, loss: 2.313353070348967e-05
step: 170, loss: 0.00023615884128957987
step: 180, loss: 2.9778151656500995e-05
step: 190, loss: 0.0001656138920225203
step: 200, loss: 1.5433530279551633e-05
step: 210, loss: 1.3548653441830538e-05
step: 220, loss: 1.5217504369502421e-05
step: 230, loss: 0.007894783280789852
step: 240, loss: 7.568171713501215e-05
step: 250, loss: 0.00016499566845595837
step: 260, loss: 0.0005697602173313498
step: 270, loss: 3.043401738977991e-05
step: 280, loss: 4.3458305299282074e-05
step: 290, loss: 4.9272377509623766e-05
step: 300, loss: 2.743608820310328e-05
step: 310, loss: 1.8160242689191364e-05
step: 320, loss: 5.999749555485323e-05
step: 330, loss: 2.1628457034239545e-05
step: 340, loss: 0.00047295616241171956
step: 350, loss: 0.00015836083912290633

```

```

step: 360, loss: 0.09381477534770966
step: 370, loss: 0.00010900741472141817
step: 380, loss: 1.8368518794886768e-05
step: 390, loss: 1.1030446330551058e-05
step: 400, loss: 0.0001025401143124327
step: 410, loss: 2.890775795094669e-05
step: 420, loss: 1.0728734196163714e-05
step: 430, loss: 1.7210315490956418e-05
step: 440, loss: 9.42490078159608e-06
step: 450, loss: 2.8392243621055968e-05
step: 460, loss: 3.681091402540915e-05
step: 470, loss: 0.0001920507347676903
step: 480, loss: 8.352034456038382e-06
step: 490, loss: 2.027947994065471e-05
step: 500, loss: 3.285522689111531e-05
step: 510, loss: 5.7484103308524936e-05
step: 520, loss: 9.953874723578338e-06
step: 530, loss: 0.00011590222857194021
/usr/local/lib/python3.7/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=0.9553488372093023, f1=0.9474662947466295,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.3929379899054766e-05
step: 10, loss: 2.4618520910735242e-05
step: 20, loss: 1.650982812861912e-05
step: 30, loss: 0.00012487084313761443
step: 40, loss: 7.668029866181314e-05
step: 50, loss: 0.0002801203227136284
step: 60, loss: 4.2740961362142116e-05
step: 70, loss: 0.0007874454604461789
step: 80, loss: 0.00018125746282748878
step: 90, loss: 1.0054433005279861e-05
step: 100, loss: 1.0520115210965741e-05
step: 110, loss: 0.00040486318175680935
step: 120, loss: 4.7378314775414765e-05
step: 130, loss: 2.503439645806793e-05
step: 140, loss: 0.00028800935251638293
step: 150, loss: 9.009848872665316e-05
step: 160, loss: 1.1488657037261873e-05
step: 170, loss: 0.011148577556014061
step: 180, loss: 1.1522164641064592e-05
step: 190, loss: 9.32059447222855e-06
step: 200, loss: 1.6308824342559092e-05
step: 210, loss: 4.338822327554226e-05

```

```

step: 220, loss: 8.642602551844902e-06
step: 230, loss: 1.0389729141024873e-05
step: 240, loss: 8.549707126803696e-05
step: 250, loss: 0.00017800097703002393
step: 260, loss: 2.2491834897664376e-05
step: 270, loss: 0.0003941576578654349
step: 280, loss: 9.49193326960085e-06
step: 290, loss: 3.2753570849308744e-05
step: 300, loss: 3.229038338758983e-05
step: 310, loss: 0.005367363803088665
step: 320, loss: 1.1801598702732008e-05
step: 330, loss: 1.7265851056436077e-05
step: 340, loss: 0.0006191634456627071
step: 350, loss: 9.205105925502721e-06
step: 360, loss: 6.0882979596499354e-05
step: 370, loss: 0.0014051805483177304
step: 380, loss: 0.0002958663390018046
step: 390, loss: 0.0020041409879922867
step: 400, loss: 0.004976102150976658
step: 410, loss: 7.4579788815754e-06
step: 420, loss: 9.53663584368769e-06
step: 430, loss: 0.0001206132656079717
step: 440, loss: 1.23156623885734e-05
step: 450, loss: 1.5801997506059706e-05
step: 460, loss: 0.010386236943304539
step: 470, loss: 1.2218780284456443e-05
step: 480, loss: 0.00022456730948761106
step: 490, loss: 1.060579961631447e-05
step: 500, loss: 0.0003154402947984636
step: 510, loss: 0.03193410485982895
step: 520, loss: 7.308361819013953e-05
step: 530, loss: 9.966228390112519e-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=0.9579831932773109, f1=0.9480580252690688,
best_f1=0.9505135387488329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005990886711515486
step: 10, loss: 0.00029757499578408897
step: 20, loss: 4.5275435695657507e-05
step: 30, loss: 0.002048956463113427
step: 40, loss: 1.1257712685619481e-05
step: 50, loss: 0.001503512030467391
step: 60, loss: 7.327784260269254e-05
step: 70, loss: 9.246078661817592e-06

```

step: 80, loss: 5.5377087846864015e-05
step: 90, loss: 8.985328349808697e-06
step: 100, loss: 8.124704436340835e-06
step: 110, loss: 1.421540855517378e-05
step: 120, loss: 8.378106031159405e-06
step: 130, loss: 1.0236940397589933e-05
step: 140, loss: 1.9191131286788732e-05
step: 150, loss: 3.955890497309156e-05
step: 160, loss: 1.0546118573984131e-05
step: 170, loss: 6.485691301350016e-06
step: 180, loss: 0.00015210133278742433
step: 190, loss: 0.0011343422811478376
step: 200, loss: 0.00020725566719193012
step: 210, loss: 8.71008014655672e-05
step: 220, loss: 1.1056499715778045e-05
step: 230, loss: 1.346668886981206e-05
step: 240, loss: 1.6438700185972266e-05
step: 250, loss: 7.690916390856728e-05
step: 260, loss: 7.566018666693708e-06
step: 270, loss: 1.2803437130060047e-05
step: 280, loss: 8.80277730175294e-06
step: 290, loss: 1.2870544196630362e-05
step: 300, loss: 1.1950426596740726e-05
step: 310, loss: 0.031538914889097214
step: 320, loss: 1.215916381624993e-05
step: 330, loss: 1.0333864338463172e-05
step: 340, loss: 2.704014696064405e-05
step: 350, loss: 0.0024089261423796415
step: 360, loss: 1.5627125321771018e-05
step: 370, loss: 1.7202666640514508e-05
step: 380, loss: 1.195802542497404e-05
step: 390, loss: 1.0583398761809804e-05
step: 400, loss: 0.00021485338220372796
step: 410, loss: 0.00010094433673657477
step: 420, loss: 9.179024345939979e-06
step: 430, loss: 6.057293376215966e-06
step: 440, loss: 2.6213874662062153e-05
step: 450, loss: 0.0027848235331475735
step: 460, loss: 0.00045232955017127097
step: 470, loss: 9.77504259935813e-06
step: 480, loss: 0.0040437583811581135
step: 490, loss: 0.0009993473067879677
step: 500, loss: 0.0011716532753780484
step: 510, loss: 7.63306252338225e-06
step: 520, loss: 1.6647147276671603e-05
step: 530, loss: 7.852851922507398e-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=0.957169459962756, f1=0.9456066945606694,
best_f1=0.9505135387488329
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 [00:35, 160.43it/s]
load_f1 = 0.9608208955223881
real_f1 = 0.9546092653252223
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, 134.36it/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: 0.4400252401828766
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=0.3023255813953489, f1=0.29268292682926833,
best_f1=0.29268292682926833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4177451431751251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.30434782608695654, f1=0.2947368421052632,
best_f1=0.2947368421052632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3947521448135376
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7878787878787878, f1=0.631578947368421,
best_f1=0.631578947368421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22839795053005219
/usr/local/lib/python3.7/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=0.7200000000000001, f1=0.7333333333333334,
best_f1=0.631578947368421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3483295738697052
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.7857142857142857, f1=0.8125000000000001,
best_f1=0.631578947368421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30713653564453125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8148148148148148, f1=0.9333333333333333,

```

```

best_f1=0.9333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2033279836177826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9655172413793104, f1=0.9032258064516129,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0316147617995739
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 8: dev_f1=0.8571428571428571, f1=0.8666666666666666,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18158762156963348
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.9032258064516129,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.066050224006176
/usr/local/lib/python3.7/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=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029049817472696304
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018314393237233162
/usr/local/lib/python3.7/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=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008068451657891273
/usr/local/lib/python3.7/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=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014318054541945457
/usr/local/lib/python3.7/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=0.9333333333333333,
best_f1=0.9032258064516129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024066617712378502
/usr/local/lib/python3.7/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=0.9333333333333333,
best_f1=0.9032258064516129
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, 122490.91it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.75
real_f1 = 0.6666666666666666
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, 134.01it/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/

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.5743722915649414

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

```

step: 10, loss: 0.44710007309913635
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.5361607670783997
step: 30, loss: 0.30558186769485474
step: 40, loss: 0.36701855063438416
step: 50, loss: 0.6437699198722839
step: 60, loss: 0.4747588336467743
step: 70, loss: 0.17032141983509064
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: 0.3467036485671997
step: 90, loss: 0.342019259929657
step: 100, loss: 0.08042691648006439
step: 110, loss: 0.07780756801366806
step: 120, loss: 0.008264567703008652
step: 130, loss: 0.013295586220920086
step: 140, loss: 0.021625198423862457
step: 150, loss: 0.07564540207386017
step: 160, loss: 0.005578198004513979
step: 170, loss: 0.07630617916584015
step: 180, loss: 0.032310571521520615
step: 190, loss: 0.18370935320854187
step: 200, loss: 0.05010583624243736
step: 210, loss: 0.13720330595970154
step: 220, loss: 0.036984123289585114
step: 230, loss: 0.0022755705285817385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9799107142857142, f1=0.9720670391061451,
best_f1=0.9720670391061451
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023418667260557413
step: 10, loss: 0.04722387716174126
step: 20, loss: 0.01881817728281021
step: 30, loss: 0.026893245056271553
step: 40, loss: 0.06767164915800095
step: 50, loss: 0.0024867418687790632
step: 60, loss: 0.0024989056400954723
step: 70, loss: 0.008926930837333202
step: 80, loss: 0.003679862944409251
step: 90, loss: 0.048735834658145905

```



```

step: 100, loss: 0.005526956170797348
step: 110, loss: 0.02963375113904476
step: 120, loss: 0.004179616924375296
step: 130, loss: 0.007029620464891195
step: 140, loss: 0.0015970952808856964
step: 150, loss: 0.14365440607070923
step: 160, loss: 0.0035668574273586273
step: 170, loss: 0.023438608273863792
step: 180, loss: 0.004347353707998991
step: 190, loss: 0.15951548516750336
step: 200, loss: 0.004064060747623444
step: 210, loss: 0.07656589150428772
step: 220, loss: 0.004672586917877197
step: 230, loss: 0.001979965716600418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842696629213483, f1=0.9818594104308391,
best_f1=0.9818594104308391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004138421267271042
step: 10, loss: 0.009773715399205685
step: 20, loss: 0.10867258161306381
step: 30, loss: 0.0009818869875743985
step: 40, loss: 0.08309037238359451
step: 50, loss: 0.05032452940940857
step: 60, loss: 0.005068259313702583
step: 70, loss: 0.0013362782774493098
step: 80, loss: 0.002162343356758356
step: 90, loss: 0.006905197165906429
step: 100, loss: 0.001208919333294034
step: 110, loss: 0.003777206176891923
step: 120, loss: 0.00029168312903493643
step: 130, loss: 0.017896942794322968
step: 140, loss: 0.007015051320195198
step: 150, loss: 0.019016696140170097
step: 160, loss: 0.0020414036698639393
step: 170, loss: 0.0029883794486522675
step: 180, loss: 0.017524920403957367
step: 190, loss: 0.031000716611742973
step: 200, loss: 0.020609546452760696
step: 210, loss: 0.002831597812473774

```

```

step: 220, loss: 0.00786617211997509
step: 230, loss: 0.02741321362555027
/usr/local/lib/python3.7/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=0.9842342342342343, f1=0.9819413092550789,
best_f1=0.9818594104308391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01860738731920719
step: 10, loss: 0.036271873861551285
step: 20, loss: 0.029918191954493523
step: 30, loss: 0.000687826017383486
step: 40, loss: 0.016050629317760468
step: 50, loss: 0.01975882798433304
step: 60, loss: 0.0028778407722711563
step: 70, loss: 0.010653626173734665
step: 80, loss: 0.0009484387701377273
step: 90, loss: 0.002484936034306884
step: 100, loss: 0.00470241904258728
step: 110, loss: 0.012104803696274757
step: 120, loss: 0.08392488956451416
step: 130, loss: 0.01673435978591442
step: 140, loss: 0.004203494172543287
step: 150, loss: 0.1267499327659607
step: 160, loss: 0.003739472944289446
step: 170, loss: 0.007216881029307842
step: 180, loss: 0.08350364118814468
step: 190, loss: 0.0009054983966052532
step: 200, loss: 0.007586784660816193
step: 210, loss: 0.005265275482088327
step: 220, loss: 0.004726800136268139
step: 230, loss: 0.002960028825327754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9864864864864865, f1=0.9863945578231292,
best_f1=0.9863945578231292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018376479856669903
step: 10, loss: 0.0026185214519500732

```

```

step: 20, loss: 0.10653875023126602
step: 30, loss: 0.0009492118842899799
step: 40, loss: 0.05654948204755783
step: 50, loss: 0.0013698870316147804
step: 60, loss: 0.003474616911262274
step: 70, loss: 0.007679760921746492
step: 80, loss: 0.08889773488044739
step: 90, loss: 0.07012929767370224
step: 100, loss: 0.0007652463973499835
step: 110, loss: 0.012406117282807827
step: 120, loss: 0.013479650020599365
step: 130, loss: 0.0007796788704581559
step: 140, loss: 0.008024466224014759
step: 150, loss: 0.007986328564584255
step: 160, loss: 0.0004555144696496427
step: 170, loss: 0.0060922312550246716
step: 180, loss: 0.002314428798854351
step: 190, loss: 0.20197270810604095
step: 200, loss: 0.024983927607536316
step: 210, loss: 0.029020365327596664
step: 220, loss: 0.0006321720429696143
step: 230, loss: 0.01278486754745245
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898762654668166, f1=0.9898305084745763,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003407563781365752
step: 10, loss: 0.011941341683268547
step: 20, loss: 0.0007652318454347551
step: 30, loss: 0.0005156206898391247
step: 40, loss: 0.0003910208761226386
step: 50, loss: 0.006102681625634432
step: 60, loss: 0.010254906490445137
step: 70, loss: 0.02119123376905918
step: 80, loss: 0.0063690682873129845
step: 90, loss: 0.11690985411405563
step: 100, loss: 0.0042825788259506226
step: 110, loss: 0.0624440535902977
step: 120, loss: 0.0004920106148347259
step: 130, loss: 0.0004364759661257267

```

```

step: 140, loss: 0.0005897974479012191
step: 150, loss: 0.00022909320250619203
step: 160, loss: 0.013948271982371807
step: 170, loss: 0.0003073598782066256
step: 180, loss: 0.0004638846730813384
step: 190, loss: 0.0005146475159563124
step: 200, loss: 0.006707819178700447
step: 210, loss: 0.005348194390535355
step: 220, loss: 0.035320453345775604
step: 230, loss: 0.0017171163344755769
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9864559819413092,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002124295337125659
step: 10, loss: 0.0009525905479677022
step: 20, loss: 0.0008593803504481912
step: 30, loss: 0.0003242132079321891
step: 40, loss: 0.0010572653263807297
step: 50, loss: 0.0008325072121806443
step: 60, loss: 0.00046426686458289623
step: 70, loss: 0.0010539134964346886
step: 80, loss: 0.00044376589357852936
step: 90, loss: 0.009943410754203796
step: 100, loss: 0.00033564839395694435
step: 110, loss: 0.0003141418274026364
step: 120, loss: 0.0005931634805165231
step: 130, loss: 0.0009005352621898055
step: 140, loss: 0.000309137802105397
step: 150, loss: 0.0024888552725315094
step: 160, loss: 0.0028324637096375227
step: 170, loss: 0.03250987455248833
step: 180, loss: 0.000201440037926659
step: 190, loss: 0.000655501673463732
step: 200, loss: 0.0030500562861561775
step: 210, loss: 0.007287728134542704
step: 220, loss: 0.00032131344778463244
step: 230, loss: 0.001113386475481093
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9886877828054299,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006995089934207499
step: 10, loss: 0.009456075727939606
step: 20, loss: 0.002401835285127163
step: 30, loss: 0.001724593574181199
step: 40, loss: 0.0013931917492300272
step: 50, loss: 0.0004784336779266596
step: 60, loss: 0.0009561444167047739
step: 70, loss: 0.00015399135008919984
step: 80, loss: 0.0050918362103402615
step: 90, loss: 0.0007678808178752661
step: 100, loss: 0.000631871516816318
step: 110, loss: 0.028416981920599937
step: 120, loss: 0.0006197458715178072
step: 130, loss: 0.004221429117023945
step: 140, loss: 0.0006872317753732204
step: 150, loss: 0.15022774040699005
step: 160, loss: 0.00767061160877347
step: 170, loss: 0.00502760661765933
step: 180, loss: 0.0016787877539172769
step: 190, loss: 0.01914018951356411
step: 200, loss: 0.011351363733410835
step: 210, loss: 0.0064377980306744576
step: 220, loss: 0.03275704011321068
step: 230, loss: 0.0006413792725652456
/usr/local/lib/python3.7/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=0.9887892376681614, f1=0.9863945578231292,
best_f1=0.9898305084745763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005184789188206196
step: 10, loss: 0.0007819424499757588
step: 20, loss: 0.0012160103069618344
step: 30, loss: 0.001286331214942038
step: 40, loss: 0.0004802948678843677
step: 50, loss: 0.00156704883556813
step: 60, loss: 0.00047247670590877533
step: 70, loss: 0.03408406302332878
step: 80, loss: 0.0016669008182361722
step: 90, loss: 0.026058971881866455
step: 100, loss: 0.0005948944599367678
step: 110, loss: 0.002183297649025917
step: 120, loss: 0.015691017732024193
step: 130, loss: 0.013480713590979576

```

```

step: 140, loss: 0.0006498204893432558
step: 150, loss: 0.001452855532988906
step: 160, loss: 0.0012460085563361645
step: 170, loss: 0.0003177584439981729
step: 180, loss: 0.0011682373005896807
step: 190, loss: 0.00010916780593106523
step: 200, loss: 0.0004929792485199869
step: 210, loss: 0.0031808754429221153
step: 220, loss: 0.0005799955106340349
step: 230, loss: 0.00019743930897675455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898989898989898, f1=0.9875424688561721,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002985650789923966
step: 10, loss: 0.000558092026039958
step: 20, loss: 0.0010136173805221915
step: 30, loss: 0.007671530824154615
step: 40, loss: 0.002124100923538208
step: 50, loss: 0.0002914874639827758
step: 60, loss: 0.0005834389012306929
step: 70, loss: 0.0024746935814619064
step: 80, loss: 0.0006908139330334961
step: 90, loss: 0.0012793304631486535
step: 100, loss: 0.0007650954648852348
step: 110, loss: 0.0020640448201447725
step: 120, loss: 0.00034965621307492256
step: 130, loss: 0.0012002303265035152
step: 140, loss: 0.0006133823771961033
step: 150, loss: 0.003977518528699875
step: 160, loss: 0.0001335281558567658
step: 170, loss: 0.0002458053349982947
step: 180, loss: 0.011420678347349167
step: 190, loss: 0.0008371495641767979
step: 200, loss: 0.0014842638047412038
step: 210, loss: 0.028584249317646027
step: 220, loss: 0.009716595523059368
step: 230, loss: 0.00032982067205011845
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.987598647125141,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002566450566519052
step: 10, loss: 0.000906776578631252
step: 20, loss: 0.0005067154997959733
step: 30, loss: 0.0003288076841272414
step: 40, loss: 0.00012252053420525044
step: 50, loss: 0.0003208410053048283
step: 60, loss: 0.006600609049201012
step: 70, loss: 0.0006064267363399267
step: 80, loss: 0.0037502788472920656
step: 90, loss: 0.21098358929157257
step: 100, loss: 0.0021133837290108204
step: 110, loss: 0.0006170551641844213
step: 120, loss: 0.0004325266636442393
step: 130, loss: 0.00401302007958293
step: 140, loss: 0.005535471718758345
step: 150, loss: 0.0034875746350735426
step: 160, loss: 0.0013377834111452103
step: 170, loss: 0.0006975557771511376
step: 180, loss: 0.0008608424104750156
step: 190, loss: 0.0007321991724893451
step: 200, loss: 0.008418493904173374
step: 210, loss: 0.0007776031852699816
step: 220, loss: 0.003161877626553178
step: 230, loss: 0.00045166665222495794
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.9864253393665158,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025391571689397097
step: 10, loss: 0.00029988473397679627
step: 20, loss: 0.021219894289970398
step: 30, loss: 0.03538540005683899
step: 40, loss: 0.001837871503084898
step: 50, loss: 0.003064072225242853
step: 60, loss: 0.0010332915699109435
step: 70, loss: 0.0005123044247739017
step: 80, loss: 0.0002447162114549428
step: 90, loss: 0.012237682938575745

```

```

step: 100, loss: 0.0003715106286108494
step: 110, loss: 0.00028680890682153404
step: 120, loss: 0.0003479073930066079
step: 130, loss: 0.00045563510502688587
step: 140, loss: 0.000980556127615273
step: 150, loss: 0.0007421004120260477
step: 160, loss: 0.00046303868293762207
step: 170, loss: 0.0009560872567817569
step: 180, loss: 0.0006083236658014357
step: 190, loss: 0.0020993428770452738
step: 200, loss: 0.0008607464260421693
step: 210, loss: 0.0030812080949544907
step: 220, loss: 0.0031120674684643745
step: 230, loss: 0.0016410858370363712
/usr/local/lib/python3.7/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=0.9886877828054299,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004740043543279171
step: 10, loss: 0.00047211573109962046
step: 20, loss: 0.0010447913082316518
step: 30, loss: 0.000815011328086257
step: 40, loss: 0.0014975781086832285
step: 50, loss: 0.0035217469558119774
step: 60, loss: 0.0011744749499484897
step: 70, loss: 0.0012943199835717678
step: 80, loss: 0.0007455253507941961
step: 90, loss: 0.0008745536324568093
step: 100, loss: 0.0013767438940703869
step: 110, loss: 0.0015591230476275086
step: 120, loss: 0.0013544118264690042
step: 130, loss: 0.0009342402336187661
step: 140, loss: 0.0004836417792830616
step: 150, loss: 0.01040409505367279
step: 160, loss: 0.010283640585839748
step: 170, loss: 0.0016216350486502051
step: 180, loss: 0.022543687373399734
step: 190, loss: 0.001144678215496242
step: 200, loss: 0.0002539900306146592
step: 210, loss: 0.000898034661076963
step: 220, loss: 0.0007074377499520779
step: 230, loss: 0.0006762165576219559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.988814317673378, f1=0.9887892376681614,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005333709996193647
step: 10, loss: 0.0003892649838235229
step: 20, loss: 0.01983489841222763
step: 30, loss: 0.0008136768592521548
step: 40, loss: 0.0005429111188277602
step: 50, loss: 0.00030388147570192814
step: 60, loss: 0.00044938037171959877
step: 70, loss: 0.0006506410427391529
step: 80, loss: 0.0006484671612270176
step: 90, loss: 0.0018053064122796059
step: 100, loss: 0.0006559669855050743
step: 110, loss: 0.0014790374552831054
step: 120, loss: 0.00024550268426537514
step: 130, loss: 0.0009665320394560695
step: 140, loss: 0.001063755014911294
step: 150, loss: 0.00012920121662318707
step: 160, loss: 0.0026309476234018803
step: 170, loss: 0.0007629484171047807
step: 180, loss: 0.0003981055924668908
step: 190, loss: 0.0022363767493516207
step: 200, loss: 0.0018171229166910052
step: 210, loss: 0.0004549737786874175
step: 220, loss: 0.0006051863892935216
step: 230, loss: 0.00028810329968109727
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9898989898989898, f1=0.9898305084745763,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05077802762389183
step: 10, loss: 0.000466290395706892
step: 20, loss: 0.0013711911160498857
step: 30, loss: 0.0006954140262678266
step: 40, loss: 0.0005100068519823253
step: 50, loss: 0.0005336144240573049
step: 60, loss: 0.03195095434784889
step: 70, loss: 0.0006934747798368335
step: 80, loss: 0.0004043809021823108
step: 90, loss: 0.00032351032132282853

```

```

step: 100, loss: 0.00029624480521306396
step: 110, loss: 0.000491702405270189
step: 120, loss: 0.08721328526735306
step: 130, loss: 0.00042370721348561347
step: 140, loss: 0.01138757448643446
step: 150, loss: 0.025318210944533348
step: 160, loss: 0.011273243464529514
step: 170, loss: 0.00026565862935967743
step: 180, loss: 0.0007001933408901095
step: 190, loss: 0.002922292100265622
step: 200, loss: 0.0005337740876711905
step: 210, loss: 0.05931911990046501
step: 220, loss: 0.001150881638750434
step: 230, loss: 0.0006321311229839921
/usr/local/lib/python3.7/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=0.9898989898989898, f1=0.9887133182844244,
best_f1=0.9875424688561721
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 [00:16, 153.08it/s]
load_f1 = 0.9876819708846584
real_f1 = 0.9799554565701558
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, 137.22it/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 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.")

step: 0, loss: 0.6158405542373657

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: 0.41711944341659546

step: 20, loss: 0.2773568630218506

step: 30, loss: 0.3308032155036926

step: 40, loss: 0.4136861264705658

step: 50, loss: 0.4394109547138214

step: 60, loss: 0.36057227849960327

step: 70, loss: 0.3176012933254242

step: 80, loss: 0.12686505913734436

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

step: 90, loss: 0.2223043292760849

step: 100, loss: 0.18170565366744995

step: 110, loss: 0.05668952316045761

step: 120, loss: 0.14985577762126923

step: 130, loss: 0.2121921330690384

step: 140, loss: 0.08863084763288498

step: 150, loss: 0.029722988605499268

step: 160, loss: 0.14825168251991272

step: 170, loss: 0.08490430563688278

step: 180, loss: 0.04614106938242912

step: 190, loss: 0.05181024596095085

```

step: 200, loss: 0.08006171882152557
step: 210, loss: 0.03554471582174301
step: 220, loss: 0.005620487499982119
step: 230, loss: 0.11760425567626953
step: 240, loss: 0.05350581556558609
step: 250, loss: 0.07133569568395615
step: 260, loss: 0.3108164668083191
step: 270, loss: 0.23510459065437317
step: 280, loss: 0.06170330196619034
step: 290, loss: 0.04356709122657776
step: 300, loss: 0.050418317317962646
step: 310, loss: 0.0745612233877182
step: 320, loss: 0.12292789667844772
step: 330, loss: 0.05942099168896675
step: 340, loss: 0.29345038533210754
step: 350, loss: 0.08447723090648651
step: 360, loss: 0.03815002739429474
step: 370, loss: 0.11391738057136536
step: 380, loss: 0.10387084633111954
step: 390, loss: 0.07845129072666168
step: 400, loss: 0.06795452535152435
step: 410, loss: 0.3082534968852997
step: 420, loss: 0.03463592752814293
step: 430, loss: 0.020868126302957535
step: 440, loss: 0.04049484059214592
step: 450, loss: 0.06088925898075104
step: 460, loss: 0.022621996700763702
step: 470, loss: 0.011389785446226597
step: 480, loss: 0.14026786386966705
step: 490, loss: 0.1460847407579422
step: 500, loss: 0.008106090128421783
step: 510, loss: 0.022719692438840866
step: 520, loss: 0.08256258070468903
step: 530, loss: 0.0032554538920521736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9462068965517241, f1=0.9447513812154696,
best_f1=0.9447513812154696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09109265357255936
step: 10, loss: 0.06670396775007248

```

step: 20, loss: 0.02029379829764366
step: 30, loss: 0.09410843998193741
step: 40, loss: 0.02790757641196251
step: 50, loss: 0.03957688808441162
step: 60, loss: 0.01605416089296341
step: 70, loss: 0.008279645815491676
step: 80, loss: 0.026868898421525955
step: 90, loss: 0.002504410920664668
step: 100, loss: 0.0522494800388813
step: 110, loss: 0.00916327815502882
step: 120, loss: 0.012337807565927505
step: 130, loss: 0.0019129383144900203
step: 140, loss: 0.057954639196395874
step: 150, loss: 0.022854022681713104
step: 160, loss: 0.04944062978029251
step: 170, loss: 0.020798305049538612
step: 180, loss: 0.0062448615208268166
step: 190, loss: 0.0036512583028525114
step: 200, loss: 0.15253695845603943
step: 210, loss: 0.03412570804357529
step: 220, loss: 0.0008395372424274683
step: 230, loss: 0.06720877438783646
step: 240, loss: 0.12398193031549454
step: 250, loss: 0.02918197028338909
step: 260, loss: 0.008110875263810158
step: 270, loss: 0.008122135885059834
step: 280, loss: 0.023547090590000153
step: 290, loss: 0.01875319518148899
step: 300, loss: 0.04544881731271744
step: 310, loss: 0.09406483173370361
step: 320, loss: 0.0701715424656868
step: 330, loss: 0.034630343317985535
step: 340, loss: 0.0328955240547657
step: 350, loss: 0.0020731580443680286
step: 360, loss: 0.09444824606180191
step: 370, loss: 0.0167030468583107
step: 380, loss: 0.1659235805273056
step: 390, loss: 0.009074116125702858
step: 400, loss: 0.024719053879380226
step: 410, loss: 0.010953557677567005
step: 420, loss: 0.02715115435421467
step: 430, loss: 0.19398146867752075
step: 440, loss: 0.00235556997358799
step: 450, loss: 0.025415748357772827
step: 460, loss: 0.06925491243600845
step: 470, loss: 0.017292121425271034
step: 480, loss: 0.026920216158032417
step: 490, loss: 0.03158726543188095

```

step: 500, loss: 0.009256678633391857
step: 510, loss: 0.005209083203226328
step: 520, loss: 0.31894734501838684
step: 530, loss: 0.09267730265855789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9503940658321743, f1=0.9517177344475395,
best_f1=0.9517177344475395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18313825130462646
step: 10, loss: 0.03103591687977314
step: 20, loss: 0.011640304699540138
step: 30, loss: 0.059670619666576385
step: 40, loss: 0.07709240913391113
step: 50, loss: 0.15736345946788788
step: 60, loss: 0.029686609283089638
step: 70, loss: 0.007787871174514294
step: 80, loss: 0.06931009143590927
step: 90, loss: 0.007530783768743277
step: 100, loss: 0.027618765830993652
step: 110, loss: 0.0418536476790905
step: 120, loss: 0.22911378741264343
step: 130, loss: 0.1157403364777565
step: 140, loss: 0.007141594309359789
step: 150, loss: 0.016958216205239296
step: 160, loss: 0.03956529125571251
step: 170, loss: 0.027481231838464737
step: 180, loss: 0.018222615122795105
step: 190, loss: 0.0033316779881715775
step: 200, loss: 0.015913672745227814
step: 210, loss: 0.012555909343063831
step: 220, loss: 0.07586969435214996
step: 230, loss: 0.03850311040878296
step: 240, loss: 0.11167450249195099
step: 250, loss: 0.08118938654661179
step: 260, loss: 0.14342868328094482
step: 270, loss: 0.0033625115174800158
step: 280, loss: 0.02899976819753647
step: 290, loss: 0.025234848260879517
step: 300, loss: 0.051491983234882355
step: 310, loss: 0.014452728442847729

```

```

step: 320, loss: 0.0578937903046608
step: 330, loss: 0.0016162883257493377
step: 340, loss: 0.004903124645352364
step: 350, loss: 0.16702914237976074
step: 360, loss: 0.04672227427363396
step: 370, loss: 0.0196770578622818
step: 380, loss: 0.003395240753889084
step: 390, loss: 0.032815635204315186
step: 400, loss: 0.040130797773599625
step: 410, loss: 0.08292258530855179
step: 420, loss: 0.0039100246503949165
step: 430, loss: 0.020280100405216217
step: 440, loss: 0.1924133449792862
step: 450, loss: 0.05383291840553284
step: 460, loss: 0.22336538136005402
step: 470, loss: 0.011747785843908787
step: 480, loss: 0.008522426709532738
step: 490, loss: 0.031322065740823746
step: 500, loss: 0.08656073361635208
step: 510, loss: 0.10283471643924713
step: 520, loss: 0.007750384975224733
step: 530, loss: 0.012941272929310799
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9523809523809524, f1=0.9522471910112359,
best_f1=0.9522471910112359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010064058005809784
step: 10, loss: 0.0011541382409632206
step: 20, loss: 0.18898825347423553
step: 30, loss: 0.10440442711114883
step: 40, loss: 0.032130781561136246
step: 50, loss: 0.05734655261039734
step: 60, loss: 0.0063822609372437
step: 70, loss: 0.01112109050154686
step: 80, loss: 0.003078785026445985
step: 90, loss: 0.03779729828238487
step: 100, loss: 0.003745466936379671
step: 110, loss: 0.08487288653850555
step: 120, loss: 0.0021326588466763496
step: 130, loss: 0.10988833755254745

```



```

step: 140, loss: 0.017949258908629417
step: 150, loss: 0.005763978231698275
step: 160, loss: 0.004098978359252214
step: 170, loss: 0.023977013304829597
step: 180, loss: 0.1345532238483429
step: 190, loss: 0.06281116604804993
step: 200, loss: 0.07075012475252151
step: 210, loss: 0.0006758155068382621
step: 220, loss: 0.0019213303457945585
step: 230, loss: 0.011021474376320839
step: 240, loss: 0.03770262748003006
step: 250, loss: 0.09075764566659927
step: 260, loss: 0.00060617970302701
step: 270, loss: 0.01731306128203869
step: 280, loss: 0.006256476975977421
step: 290, loss: 0.07162892818450928
step: 300, loss: 0.0019282522844150662
step: 310, loss: 0.005533857736736536
step: 320, loss: 0.08928599208593369
step: 330, loss: 0.0961974635720253
step: 340, loss: 0.014128240756690502
step: 350, loss: 0.12597113847732544
step: 360, loss: 0.02370181307196617
step: 370, loss: 0.0020515990909188986
step: 380, loss: 0.07092610001564026
step: 390, loss: 0.004008039366453886
step: 400, loss: 0.08064645528793335
step: 410, loss: 0.004924590699374676
step: 420, loss: 0.014720306731760502
step: 430, loss: 0.03511922061443329
step: 440, loss: 0.027434129267930984
step: 450, loss: 0.05756847932934761
step: 460, loss: 0.026734447106719017
step: 470, loss: 0.009721084497869015
step: 480, loss: 0.0013643059646710753
step: 490, loss: 0.015440410003066063
step: 500, loss: 0.05767488479614258
step: 510, loss: 0.08354293555021286
step: 520, loss: 0.00866856798529625
step: 530, loss: 0.11980301141738892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.954001839926403, f1=0.950668510834486,

```

```
best_f1=0.950668510834486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033454972435720265
step: 10, loss: 0.007509918883442879
step: 20, loss: 0.0006146585801616311
step: 30, loss: 0.014241800643503666
step: 40, loss: 0.008343391120433807
step: 50, loss: 0.03387153148651123
step: 60, loss: 0.020598003640770912
step: 70, loss: 0.012451332993805408
step: 80, loss: 0.0023575129453092813
step: 90, loss: 0.08543536067008972
step: 100, loss: 0.1463022381067276
step: 110, loss: 0.007958617992699146
step: 120, loss: 0.10838716477155685
step: 130, loss: 0.023533830419182777
step: 140, loss: 0.014262745156884193
step: 150, loss: 0.017363442108035088
step: 160, loss: 0.015757476910948753
step: 170, loss: 0.024475833401083946
step: 180, loss: 0.004048509057611227
step: 190, loss: 0.0011513297213241458
step: 200, loss: 0.005083238705992699
step: 210, loss: 0.01756351627409458
step: 220, loss: 0.010301653295755386
step: 230, loss: 0.0009616132592782378
step: 240, loss: 0.02336747571825981
step: 250, loss: 0.09888792037963867
step: 260, loss: 0.0024494300596415997
step: 270, loss: 0.03693588823080063
step: 280, loss: 0.02017081342637539
step: 290, loss: 0.0005039595998823643
step: 300, loss: 0.058782465755939484
step: 310, loss: 0.015556100755929947
step: 320, loss: 0.008049601688981056
step: 330, loss: 0.002061843406409025
step: 340, loss: 0.027248786762356758
step: 350, loss: 0.0010339203290641308
step: 360, loss: 9.550303366268054e-05
step: 370, loss: 0.00040621653897687793
step: 380, loss: 0.0016876982990652323
step: 390, loss: 0.05689619109034538
step: 400, loss: 0.004217368084937334
step: 410, loss: 0.06994103640317917
step: 420, loss: 0.19406382739543915
step: 430, loss: 0.008678801357746124
```

```
step: 440, loss: 0.0004749662766698748
step: 450, loss: 0.00858917273581028
step: 460, loss: 0.07911408692598343
step: 470, loss: 0.10325106233358383
step: 480, loss: 0.2506561279296875
step: 490, loss: 0.018347160890698433
step: 500, loss: 0.030724141746759415
step: 510, loss: 0.15659204125404358
step: 520, loss: 0.14829425513744354
step: 530, loss: 0.01845848746597767
/usr/local/lib/python3.7/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=0.9534883720930233, f1=0.9502556950255696,
best_f1=0.950668510834486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01676454395055771
step: 10, loss: 0.0006936874706298113
step: 20, loss: 0.0027526712510734797
step: 30, loss: 0.000511283113155514
step: 40, loss: 0.00024053192464634776
step: 50, loss: 0.005994661245495081
step: 60, loss: 0.005058398470282555
step: 70, loss: 0.0002133187954314053
step: 80, loss: 0.0003820249403361231
step: 90, loss: 0.008062900975346565
step: 100, loss: 0.0012727777939289808
step: 110, loss: 0.0015023587038740516
step: 120, loss: 0.018252624198794365
step: 130, loss: 0.005442756228148937
step: 140, loss: 0.002924702363088727
step: 150, loss: 0.00015281958621926606
step: 160, loss: 0.014781322330236435
step: 170, loss: 0.00011882163380505517
step: 180, loss: 0.00032186813768930733
step: 190, loss: 0.07629508525133133
step: 200, loss: 0.03253428265452385
step: 210, loss: 0.0012355622602626681
step: 220, loss: 0.0007638218812644482
step: 230, loss: 0.040402576327323914
step: 240, loss: 0.07144582271575928
step: 250, loss: 0.013399794697761536
step: 260, loss: 0.0045477598905563354
step: 270, loss: 0.016437970101833344
step: 280, loss: 0.011302761733531952
step: 290, loss: 0.02499813586473465
```

```

step: 300, loss: 0.05084533616900444
step: 310, loss: 0.06466232240200043
step: 320, loss: 0.01797865703701973
step: 330, loss: 0.005254116840660572
step: 340, loss: 0.0002757288748398423
step: 350, loss: 0.0018246854888275266
step: 360, loss: 0.0028406509663909674
step: 370, loss: 0.008239013142883778
step: 380, loss: 0.033891137689352036
step: 390, loss: 0.0005040079704485834
step: 400, loss: 0.035180285573005676
step: 410, loss: 0.00044351298129186034
step: 420, loss: 0.011657197028398514
step: 430, loss: 0.004686822649091482
step: 440, loss: 0.0006451415247283876
step: 450, loss: 0.1545322835445404
step: 460, loss: 0.004722157493233681
step: 470, loss: 0.007569117471575737
step: 480, loss: 0.0011077558156102896
step: 490, loss: 0.006887447088956833
step: 500, loss: 0.0005162869929336011
step: 510, loss: 0.1117333248257637
step: 520, loss: 0.001078820088878274
step: 530, loss: 0.00994635559618473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9557522123893806, f1=0.9533582089552238,
best_f1=0.9533582089552238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031940967310220003
step: 10, loss: 0.0016320161521434784
step: 20, loss: 0.0061777615919709206
step: 30, loss: 0.025112245231866837
step: 40, loss: 0.001147128758020699
step: 50, loss: 0.05883709713816643
step: 60, loss: 0.022508373484015465
step: 70, loss: 0.0010138169163838029
step: 80, loss: 0.00019315729150548577
step: 90, loss: 0.001656481297686696
step: 100, loss: 0.0018104768823832273
step: 110, loss: 0.0002519839908927679

```

```

step: 120, loss: 0.005713006481528282
step: 130, loss: 0.010903499089181423
step: 140, loss: 0.02160070464015007
step: 150, loss: 0.004004968795925379
step: 160, loss: 0.0002305870730197057
step: 170, loss: 0.023667875677347183
step: 180, loss: 0.0766403079032898
step: 190, loss: 0.03110423870384693
step: 200, loss: 0.002291819080710411
step: 210, loss: 0.0020030003506690264
step: 220, loss: 0.04360814392566681
step: 230, loss: 0.030380744487047195
step: 240, loss: 0.0006627212278544903
step: 250, loss: 0.07710957527160645
step: 260, loss: 0.0019587520509958267
step: 270, loss: 0.02181209810078144
step: 280, loss: 0.0012882882729172707
step: 290, loss: 0.037124209105968475
step: 300, loss: 0.00099824252538383
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 310, loss: 0.002095813164487481
step: 320, loss: 0.0003457093262113631
step: 330, loss: 0.0004891860298812389
step: 340, loss: 0.0022151281591504812
step: 350, loss: 0.00679879542440176
step: 360, loss: 0.036730244755744934
step: 370, loss: 0.004278767388314009
step: 380, loss: 0.00037436577258631587
step: 390, loss: 0.007211730815470219
step: 400, loss: 0.011095840483903885
step: 410, loss: 0.006270872429013252
step: 420, loss: 0.03716619312763214
step: 430, loss: 0.010855036787688732
step: 440, loss: 0.001222822698764503
step: 450, loss: 0.008087821304798126
step: 460, loss: 0.035014811903238297
step: 470, loss: 0.14681582152843475
step: 480, loss: 0.0016967556439340115
step: 490, loss: 0.016967693343758583
step: 500, loss: 0.0033960314467549324
step: 510, loss: 0.0001677085820119828
step: 520, loss: 0.0031471869442611933
step: 530, loss: 0.0003524304192978889
/usr/local/lib/python3.7/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=0.9557195571955718, f1=0.9497000461467465,
best_f1=0.9533582089552238

```

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

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

```
step: 0, loss: 0.00023769194376654923  
step: 10, loss: 0.011089185252785683  
step: 20, loss: 0.019134540110826492  
step: 30, loss: 0.0015694205649197102  
step: 40, loss: 5.88174480071757e-05  
step: 50, loss: 0.0009652064763940871  
step: 60, loss: 0.00016520269855391234  
step: 70, loss: 0.004494276363402605  
step: 80, loss: 0.0022957504261285067  
step: 90, loss: 0.0004342089523561299  
step: 100, loss: 0.0028243751730769873  
step: 110, loss: 0.0013227362651377916  
step: 120, loss: 0.003236201126128435  
step: 130, loss: 0.0008335710153914988  
step: 140, loss: 0.002582845278084278  
step: 150, loss: 0.001846443279646337  
step: 160, loss: 0.0002468322345521301  
step: 170, loss: 0.0637805238366127  
step: 180, loss: 0.004856647923588753  
step: 190, loss: 0.006551705300807953  
step: 200, loss: 0.015426578931510448  
step: 210, loss: 0.07736058533191681  
step: 220, loss: 0.0022976784966886044  
step: 230, loss: 0.10906940698623657  
step: 240, loss: 0.006787484977394342  
step: 250, loss: 0.0017000045627355576  
step: 260, loss: 0.0019061560742557049  
step: 270, loss: 0.003175922203809023  
step: 280, loss: 0.00016853350098244846  
step: 290, loss: 0.01564699225127697  
step: 300, loss: 0.00042179712909273803  
step: 310, loss: 0.0004953923635184765  
step: 320, loss: 0.0007283922750502825  
step: 330, loss: 0.001037173205986619  
step: 340, loss: 0.04313024878501892  
step: 350, loss: 0.0013791966484859586  
step: 360, loss: 0.042086634784936905  
step: 370, loss: 0.04452044516801834  
step: 380, loss: 0.0007223320426419377  
step: 390, loss: 0.005623293109238148  
step: 400, loss: 0.001984128262847662  
step: 410, loss: 6.933951226528734e-05  
step: 420, loss: 0.0004304377071093768  
step: 430, loss: 0.01119744498282671  
step: 440, loss: 0.00401519937440753
```

```
step: 450, loss: 0.0006960108876228333
step: 460, loss: 0.0064778681844472885
step: 470, loss: 0.053742483258247375
step: 480, loss: 0.0072577777318656445
step: 490, loss: 0.024249866604804993
step: 500, loss: 0.001459192018955946
step: 510, loss: 0.007168047595769167
step: 520, loss: 0.0014386451803147793
step: 530, loss: 0.0015423569129779935
/usr/local/lib/python3.7/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=0.9543568464730291, f1=0.9509399358092617,
best_f1=0.9533582089552238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013917424948886037
step: 10, loss: 0.004375153686851263
step: 20, loss: 0.000452598265837878
step: 30, loss: 0.12848250567913055
step: 40, loss: 0.009130912832915783
step: 50, loss: 0.0005417391657829285
step: 60, loss: 0.0007846099324524403
step: 70, loss: 0.01383074652403593
step: 80, loss: 0.004304119385778904
step: 90, loss: 0.03887423872947693
step: 100, loss: 0.0004987845895811915
step: 110, loss: 0.049799297004938126
step: 120, loss: 0.01080501638352871
step: 130, loss: 0.004913197830319405
step: 140, loss: 0.006945180706679821
step: 150, loss: 0.005797653924673796
step: 160, loss: 0.005103583913296461
step: 170, loss: 0.007462484762072563
step: 180, loss: 0.000381141813704744
step: 190, loss: 0.08450475335121155
step: 200, loss: 0.0004385880893096328
step: 210, loss: 0.0010864080395549536
step: 220, loss: 0.00032689838553778827
step: 230, loss: 0.0009312012698501348
step: 240, loss: 0.0033732035662978888
step: 250, loss: 0.00024223503714893013
step: 260, loss: 0.0005017257062718272
step: 270, loss: 0.016681313514709473
step: 280, loss: 0.007483817636966705
step: 290, loss: 0.0005324056837707758
step: 300, loss: 0.0006347846938297153
```

```

step: 310, loss: 0.0750197321176529
step: 320, loss: 0.0003807605244219303
step: 330, loss: 0.0009095733403228223
step: 340, loss: 0.05750410258769989
step: 350, loss: 0.014277655631303787
step: 360, loss: 0.004833436105400324
step: 370, loss: 0.0008231776300817728
step: 380, loss: 0.00025290268240496516
step: 390, loss: 0.0013983940007165074
step: 400, loss: 0.029988804832100868
step: 410, loss: 0.0004192031337879598
step: 420, loss: 0.000794916704762727
step: 430, loss: 0.002737276488915086
step: 440, loss: 0.002060400787740946
step: 450, loss: 0.03364059701561928
step: 460, loss: 1.2218792107887566e-05
step: 470, loss: 0.00044241821160539985
step: 480, loss: 3.684313560370356e-05
step: 490, loss: 0.026046564802527428
step: 500, loss: 6.870489596622065e-05
step: 510, loss: 0.006452751345932484
step: 520, loss: 0.0036546981427818537
step: 530, loss: 0.05027452111244202
/usr/local/lib/python3.7/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=0.9535747446610957, f1=0.9507593189139439,
best_f1=0.9533582089552238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015998417511582375
step: 10, loss: 0.0057463753037154675
step: 20, loss: 0.0004958180943503976
step: 30, loss: 0.0030298687051981688
step: 40, loss: 0.013227641582489014
step: 50, loss: 0.00017867414862848818
step: 60, loss: 0.02547002211213112
step: 70, loss: 0.0003574719012249261
step: 80, loss: 0.006330670788884163
step: 90, loss: 0.021711859852075577
step: 100, loss: 0.004552703816443682
step: 110, loss: 0.006048970390111208
step: 120, loss: 0.0006948878872208297
step: 130, loss: 0.030787166208028793
step: 140, loss: 7.824433851055801e-05
step: 150, loss: 0.00043211106094531715
step: 160, loss: 0.02958489954471588

```



```

step: 170, loss: 0.001360360300168395
step: 180, loss: 0.0005912843625992537
step: 190, loss: 1.2107001566619147e-05
step: 200, loss: 0.006108431611210108
step: 210, loss: 0.009478999301791191
step: 220, loss: 0.0009979113237932324
step: 230, loss: 0.0009240766521543264
step: 240, loss: 0.002171094296500087
step: 250, loss: 0.001678145956248045
step: 260, loss: 0.009053193032741547
step: 270, loss: 6.623539229622111e-05
step: 280, loss: 0.08703659474849701
step: 290, loss: 0.0009487906936556101
step: 300, loss: 0.0035881863441318274
step: 310, loss: 0.022164883092045784
step: 320, loss: 0.0977291539311409
step: 330, loss: 0.02193455770611763
step: 340, loss: 0.0019528098637238145
step: 350, loss: 0.00028670733445324004
step: 360, loss: 0.0002188963844673708
step: 370, loss: 0.02333655208349228
step: 380, loss: 0.09376240521669388
step: 390, loss: 0.00022052603890188038
step: 400, loss: 0.0015253175515681505
step: 410, loss: 0.005036801565438509
step: 420, loss: 0.0012935823760926723
step: 430, loss: 0.0009847335750237107
step: 440, loss: 0.004613130819052458
step: 450, loss: 0.0007218327955342829
step: 460, loss: 0.00011740988702513278
step: 470, loss: 0.0037778904661536217
step: 480, loss: 0.0005066945450380445
step: 490, loss: 0.03671729937195778
step: 500, loss: 0.0462476946413517
step: 510, loss: 6.490397936431691e-05
step: 520, loss: 0.026819821447134018
step: 530, loss: 0.10107108950614929
/usr/local/lib/python3.7/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=0.9526462395543175, f1=0.9491211840888066,
best_f1=0.9533582089552238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.346038521267474e-05
step: 10, loss: 0.08870583772659302
step: 20, loss: 0.0006324139540083706

```

step: 30, loss: 0.00046157813630998135
step: 40, loss: 0.0020417231135070324
step: 50, loss: 0.00028008842491544783
step: 60, loss: 0.004696215968579054
step: 70, loss: 0.00011380224168533459
step: 80, loss: 0.0020064618438482285
step: 90, loss: 0.003642653813585639
step: 100, loss: 0.0006384194130077958
step: 110, loss: 0.0012327793519943953
step: 120, loss: 0.013041014783084393
step: 130, loss: 0.007840840145945549
step: 140, loss: 0.00025642625405453146
step: 150, loss: 0.002297211205586791
step: 160, loss: 0.0011678158771246672
step: 170, loss: 0.011553548276424408
step: 180, loss: 7.555371121270582e-05
step: 190, loss: 0.00016786769265308976
step: 200, loss: 3.322918928461149e-05
step: 210, loss: 9.69267712207511e-05
step: 220, loss: 0.03163628652691841
step: 230, loss: 0.0004249599005561322
step: 240, loss: 0.0031388713978230953
step: 250, loss: 1.0311428923159838e-05
step: 260, loss: 0.0021062283776700497
step: 270, loss: 0.0029034747276455164
step: 280, loss: 0.0006067834910936654
step: 290, loss: 0.00012658422929234803
step: 300, loss: 0.005566674750298262
step: 310, loss: 0.0019685933366417885
step: 320, loss: 0.00014260446187108755
step: 330, loss: 6.0684756135742646e-06
step: 340, loss: 0.08901235461235046
step: 350, loss: 0.0016636672662571073
step: 360, loss: 0.0002457659284118563
step: 370, loss: 0.00142505147960037
step: 380, loss: 0.00029453542083501816
step: 390, loss: 0.01221819780766964
step: 400, loss: 0.006551289465278387
step: 410, loss: 0.002243253868073225
step: 420, loss: 0.002412979258224368
step: 430, loss: 0.001629096339456737
step: 440, loss: 0.0006237755296751857
step: 450, loss: 0.0025381066370755434
step: 460, loss: 0.0015414447989314795
step: 470, loss: 0.00026735433493740857
step: 480, loss: 0.0026594598311930895
step: 490, loss: 0.0038270154036581516
step: 500, loss: 0.00037537593743763864

```

step: 510, loss: 0.010694475844502449
step: 520, loss: 0.0023383027873933315
step: 530, loss: 0.02227558009326458
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9578508568781844, f1=0.9494949494949496,
best_f1=0.9494949494949496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010617297375574708
step: 10, loss: 0.0003735973732545972
step: 20, loss: 0.0011518162209540606
step: 30, loss: 1.367501954518957e-05
step: 40, loss: 0.0027860833797603846
step: 50, loss: 0.00011698752496158704
step: 60, loss: 8.022219844860956e-05
step: 70, loss: 0.05066565051674843
step: 80, loss: 0.00031479785684496164
step: 90, loss: 0.004064904525876045
step: 100, loss: 0.1961756944656372
step: 110, loss: 0.00028562149964272976
step: 120, loss: 0.0002587856142781675
step: 130, loss: 0.0012887833872810006
step: 140, loss: 0.00012668021372519433
step: 150, loss: 0.00025293734506703913
step: 160, loss: 0.0009158664615824819
step: 170, loss: 0.0001252238726010546
step: 180, loss: 0.00018425022426526994
step: 190, loss: 0.003998785745352507
step: 200, loss: 0.0029137050732970238
step: 210, loss: 7.800567982485518e-05
step: 220, loss: 3.006269434990827e-05
step: 230, loss: 0.0001060659415088594
step: 240, loss: 0.0012691044248640537
step: 250, loss: 9.987395969801582e-06
step: 260, loss: 0.00030134458211250603
step: 270, loss: 0.0572233684360981
step: 280, loss: 5.0108479626942426e-05
step: 290, loss: 0.001551485271193087
step: 300, loss: 0.002169952495023608
step: 310, loss: 0.0001464628876419738
step: 320, loss: 4.810539758182131e-05

```

```

step: 330, loss: 0.0009314264170825481
step: 340, loss: 0.005857087206095457
step: 350, loss: 4.637592428480275e-05
step: 360, loss: 0.0001910663559101522
step: 370, loss: 0.007261720020323992
step: 380, loss: 0.0021568394731730223
step: 390, loss: 0.0008562557050026953
step: 400, loss: 2.6835927201318555e-05
step: 410, loss: 0.000785288808401674
step: 420, loss: 0.011678666807711124
step: 430, loss: 0.0003815802338067442
step: 440, loss: 0.0011372038861736655
step: 450, loss: 0.0044065737165510654
step: 460, loss: 9.808461982174776e-06
step: 470, loss: 0.04263174533843994
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 480, loss: 0.00126453407574445
step: 490, loss: 0.000957930984441191
step: 500, loss: 2.3143451471696608e-05
step: 510, loss: 0.033534321933984756
step: 520, loss: 0.003453037003055215
step: 530, loss: 0.0003799733240157366
/usr/local/lib/python3.7/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=0.9570552147239263, f1=0.9505882352941175,
best_f1=0.9494949494949496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013969045539852232
step: 10, loss: 8.62087108544074e-05
step: 20, loss: 3.005956023116596e-05
step: 30, loss: 9.652093467593659e-06
step: 40, loss: 0.005730270873755217
step: 50, loss: 0.007068990729749203
step: 60, loss: 0.0010998923098668456
step: 70, loss: 0.0035993061028420925
step: 80, loss: 0.013643953017890453
step: 90, loss: 0.0012864531017839909
step: 100, loss: 0.0004996820935048163
step: 110, loss: 0.0003581511555239558
step: 120, loss: 0.00022945000091567636
step: 130, loss: 0.0008716805023141205
step: 140, loss: 0.00459985202178359
step: 150, loss: 0.0012969112722203135
step: 160, loss: 0.002851961413398385
step: 170, loss: 0.0055154841393232346

```

```

step: 180, loss: 0.0001549065054859966
step: 190, loss: 0.0006725869607180357
step: 200, loss: 1.2431084542186e-05
step: 210, loss: 8.720826372154988e-06
step: 220, loss: 0.0003923167532775551
step: 230, loss: 0.0030278086196631193
step: 240, loss: 0.0005095288506709039
step: 250, loss: 0.0008348748087882996
step: 260, loss: 0.0005361600196920335
step: 270, loss: 0.009881886653602123
step: 280, loss: 3.952764018322341e-05
step: 290, loss: 0.04588150233030319
step: 300, loss: 1.7228672732017003e-05
step: 310, loss: 0.0011241351021453738
step: 320, loss: 1.0006019692809787e-05
step: 330, loss: 0.0005305049708113074
step: 340, loss: 0.001400992157869041
step: 350, loss: 0.00028080944321118295
step: 360, loss: 0.09006191790103912
step: 370, loss: 0.01260923221707344
step: 380, loss: 0.0004147382569499314
step: 390, loss: 0.00021239757188595831
step: 400, loss: 0.07667592167854309
step: 410, loss: 0.00018615316366776824
step: 420, loss: 0.0017782894428819418
step: 430, loss: 2.9602979338960722e-05
step: 440, loss: 2.6430454454384744e-05
step: 450, loss: 0.0008078302489593625
step: 460, loss: 0.01858857460319996
step: 470, loss: 0.0014858641661703587
step: 480, loss: 1.4159448255668394e-05
step: 490, loss: 0.00010763212776510045
step: 500, loss: 0.0011626643827185035
step: 510, loss: 5.948357647866942e-05
step: 520, loss: 0.00010715849930420518
step: 530, loss: 8.67018970893696e-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=0.9579045837231057, f1=0.9514018691588785,
best_f1=0.9514018691588785
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.000996642280369997
step: 10, loss: 0.0016344561008736491
step: 20, loss: 0.00012380207772366703
step: 30, loss: 0.04300723969936371
step: 40, loss: 0.000615505559835583
step: 50, loss: 0.0006956718862056732
step: 60, loss: 0.00017879826191347092
step: 70, loss: 0.008903379552066326
step: 80, loss: 0.00017109439068008214
step: 90, loss: 9.238638085662387e-06
step: 100, loss: 0.0001036296525853686
step: 110, loss: 0.00011734345025615767
step: 120, loss: 3.0072647859924473e-05
step: 130, loss: 0.00019635869830381125
step: 140, loss: 0.0009567553061060607
step: 150, loss: 0.0009626292739994824
step: 160, loss: 0.00015604471263941377
step: 170, loss: 0.01786733604967594
step: 180, loss: 0.017543481662869453
step: 190, loss: 7.905974052846432e-05
step: 200, loss: 0.00011310993431834504
step: 210, loss: 0.008681870996952057
step: 220, loss: 1.9902594431187026e-05
step: 230, loss: 2.387782842561137e-05
step: 240, loss: 0.0014694490237161517
step: 250, loss: 0.0005250556860119104
step: 260, loss: 7.077310146996751e-05
step: 270, loss: 0.00012200763740111142
step: 280, loss: 7.551417365903035e-05
step: 290, loss: 2.9965973226353526e-05
step: 300, loss: 0.00036311798612587154
step: 310, loss: 0.002984544960781932
step: 320, loss: 1.3570735063694883e-05
step: 330, loss: 0.0009274948388338089
step: 340, loss: 0.000226557778660208
step: 350, loss: 0.00033748295390978456
step: 360, loss: 0.00019714116933755577
step: 370, loss: 9.805364243220538e-05
step: 380, loss: 0.03897198662161827
step: 390, loss: 7.75406151660718e-05
step: 400, loss: 0.06900647282600403
step: 410, loss: 0.0005356109468266368
step: 420, loss: 6.71902162139304e-05
step: 430, loss: 0.004090990871191025
step: 440, loss: 0.03560097515583038
step: 450, loss: 0.00012344551214482635
step: 460, loss: 0.02033879980444908
step: 470, loss: 4.274133243598044e-05

```

step: 480, loss: 0.011210747063159943
step: 490, loss: 0.0013923872029408813
step: 500, loss: 0.0024044597521424294
step: 510, loss: 0.021741900593042374
step: 520, loss: 0.0006671456503681839
step: 530, loss: 0.0014126035384833813
/usr/local/lib/python3.7/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=0.9567240577012563, f1=0.95, best_f1=0.9514018691588785
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003587696934118867
step: 10, loss: 0.0024175960570573807
step: 20, loss: 0.00168482749722898
step: 30, loss: 0.00457850843667984
step: 40, loss: 2.016347207245417e-05
step: 50, loss: 0.0010283057345077395
step: 60, loss: 0.0005160585860721767
step: 70, loss: 9.052857785718516e-05
step: 80, loss: 0.00014686756185255945
step: 90, loss: 0.0004747055936604738
step: 100, loss: 6.275824853219092e-05
step: 110, loss: 0.001970703713595867
step: 120, loss: 1.8595936126075685e-05
step: 130, loss: 0.0004501736839301884
step: 140, loss: 2.705976294237189e-05
step: 150, loss: 4.107530912733637e-05
step: 160, loss: 0.0007318123825825751
step: 170, loss: 3.078689405811019e-05
step: 180, loss: 0.018846116960048676
step: 190, loss: 0.0189739391207695
step: 200, loss: 0.0010217784438282251
step: 210, loss: 2.8475147701101378e-05
step: 220, loss: 5.312259600032121e-05
step: 230, loss: 8.981043356470764e-05
step: 240, loss: 0.0003621563664637506
step: 250, loss: 1.3719927665079013e-05
step: 260, loss: 6.932727046660148e-06
step: 270, loss: 6.694303920085076e-06
step: 280, loss: 3.075789936701767e-05
step: 290, loss: 0.004908808972686529
step: 300, loss: 0.00014762043429072946
step: 310, loss: 0.034794844686985016
step: 320, loss: 2.8514055884443223e-05
step: 330, loss: 0.0009784535504877567
step: 340, loss: 0.0005253351991996169

```

```

step: 350, loss: 0.000952504517044872
step: 360, loss: 0.0004198704264126718
step: 370, loss: 0.0014234682312235236
step: 380, loss: 5.853973561897874e-05
step: 390, loss: 0.0001234908268088475
step: 400, loss: 0.04738297313451767
step: 410, loss: 5.56770246475935e-05
step: 420, loss: 0.0003526898508425802
step: 430, loss: 2.0409070202731527e-05
step: 440, loss: 0.0008346190443262458
step: 450, loss: 0.0181331317871809
step: 460, loss: 0.0017714878777042031
step: 470, loss: 2.7126434360980056e-05
step: 480, loss: 0.0019589816220104694
step: 490, loss: 0.00018828685279004276
step: 500, loss: 0.011356990784406662
step: 510, loss: 3.211684452253394e-05
step: 520, loss: 1.4204149920260534e-05
step: 530, loss: 6.604172813240439e-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=0.9576547231270358, f1=0.949560388708931,
best_f1=0.9514018691588785
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, 162.93it/s]
load_f1 = 0.9586623316302832
real_f1 = 0.958217270194986
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, 134.77it/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/
```

```
Downloading: 100% 481/481 [00:00<00:00, 410kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 793kB/s]
Downloading: 100% 456k/456k [00:01<00:00, 444kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 70.7MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

```
step: 0, loss: 0.5207846164703369
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4583336412906647

```

step: 20, loss: 0.445688396692276
step: 30, loss: 0.39327630400657654
step: 40, loss: 0.2884081304073334
step: 50, loss: 0.5252454876899719
step: 60, loss: 0.468273788690567
step: 70, loss: 0.2984233498573303
step: 80, loss: 0.3483205735683441
step: 90, loss: 0.2757798433303833
step: 100, loss: 0.27999064326286316
step: 110, loss: 0.2044616937637329
step: 120, loss: 0.33059394359588623
step: 130, loss: 0.18320536613464355
step: 140, loss: 0.4303092956542969
step: 150, loss: 0.19008630514144897
step: 160, loss: 0.3752618730068207
step: 170, loss: 0.2172130048274994
step: 180, loss: 0.25068140029907227
step: 190, loss: 0.26536089181900024
step: 200, loss: 0.18538987636566162
step: 210, loss: 0.24601605534553528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6074600355239786, f1=0.6355140186915887,
best_f1=0.6355140186915887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11538224667310715
step: 10, loss: 0.10552714765071869
step: 20, loss: 0.2779856324195862
step: 30, loss: 0.1444084346294403
step: 40, loss: 0.500154435634613
step: 50, loss: 0.09075362980365753
step: 60, loss: 0.2405068427324295
step: 70, loss: 0.194619283080101
step: 80, loss: 0.18036842346191406
step: 90, loss: 0.1729322075843811
step: 100, loss: 0.2765161097049713
step: 110, loss: 0.21875491738319397
step: 120, loss: 0.17168857157230377
step: 130, loss: 0.07336819171905518
step: 140, loss: 0.057928360998630524
step: 150, loss: 0.2827250063419342

```

```

step: 160, loss: 0.08107718825340271
step: 170, loss: 0.2642829716205597
step: 180, loss: 0.1859285682439804
step: 190, loss: 0.23291967809200287
step: 200, loss: 0.09313874691724777
step: 210, loss: 0.12135863304138184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6746031746031745, f1=0.7094188376753506,
best_f1=0.7094188376753506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08694341778755188
step: 10, loss: 0.02947109565138817
step: 20, loss: 0.25807827711105347
step: 30, loss: 0.059748224914073944
step: 40, loss: 0.20026272535324097
step: 50, loss: 0.23905128240585327
step: 60, loss: 0.20462220907211304
step: 70, loss: 0.10388270765542984
step: 80, loss: 0.14923140406608582
step: 90, loss: 0.05111650004982948
step: 100, loss: 0.15136073529720306
step: 110, loss: 0.07606470584869385
step: 120, loss: 0.13828007876873016
step: 130, loss: 0.18924914300441742
step: 140, loss: 0.083883136510849
step: 150, loss: 0.1558329313993454
step: 160, loss: 0.21929684281349182
step: 170, loss: 0.14514410495758057
step: 180, loss: 0.10160745680332184
step: 190, loss: 0.019198130816221237
step: 200, loss: 0.04375554248690605
step: 210, loss: 0.10450778156518936
/usr/local/lib/python3.7/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=0.6705882352941176, f1=0.714828897338403,
best_f1=0.7094188376753506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.055523402988910675
step: 10, loss: 0.16231413185596466
step: 20, loss: 0.08529991656541824
step: 30, loss: 0.11696624755859375
step: 40, loss: 0.0707816332578659
step: 50, loss: 0.07491540908813477
step: 60, loss: 0.17995263636112213
step: 70, loss: 0.06130669638514519
step: 80, loss: 0.04473553225398064
step: 90, loss: 0.04421816021203995
step: 100, loss: 0.17457890510559082
step: 110, loss: 0.2383926808834076
step: 120, loss: 0.1536015272140503
step: 130, loss: 0.4575226306915283
step: 140, loss: 0.23009148240089417
step: 150, loss: 0.15913943946361542
step: 160, loss: 0.11644009500741959
step: 170, loss: 0.0820484459400177
step: 180, loss: 0.0038439291529357433
step: 190, loss: 0.13248789310455322
step: 200, loss: 0.1581052988767624
step: 210, loss: 0.29171183705329895
/usr/local/lib/python3.7/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=0.6638297872340426, f1=0.6796536796536797,
best_f1=0.7094188376753506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08558028936386108
step: 10, loss: 0.033634692430496216
step: 20, loss: 0.04212816432118416
step: 30, loss: 0.015594588592648506
step: 40, loss: 0.05383118987083435
step: 50, loss: 0.06565535068511963
step: 60, loss: 0.019838426262140274
step: 70, loss: 0.1142880767583847
step: 80, loss: 0.07167309522628784
step: 90, loss: 0.08647394180297852
step: 100, loss: 0.017689961940050125
step: 110, loss: 0.11841052770614624
step: 120, loss: 0.11442255228757858
step: 130, loss: 0.04184098541736603
step: 140, loss: 0.17405885457992554
step: 150, loss: 0.07543058693408966
step: 160, loss: 0.05942443758249283
step: 170, loss: 0.0412585623562336

```

```

step: 180, loss: 0.05875193327665329
step: 190, loss: 0.2162727564573288
step: 200, loss: 0.044552870094776154
step: 210, loss: 0.19152992963790894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7118644067796611, f1=0.7446808510638298,
best_f1=0.7446808510638298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03373784199357033
step: 10, loss: 0.08161547034978867
step: 20, loss: 0.02396578900516033
step: 30, loss: 0.04842732474207878
step: 40, loss: 0.03891347721219063
step: 50, loss: 0.03425072878599167
step: 60, loss: 0.11032549291849136
step: 70, loss: 0.053147707134485245
step: 80, loss: 0.056327540427446365
step: 90, loss: 0.02265394851565361
step: 100, loss: 0.13084886968135834
step: 110, loss: 0.24397709965705872
step: 120, loss: 0.015789058059453964
step: 130, loss: 0.007238028571009636
step: 140, loss: 0.08796567469835281
step: 150, loss: 0.09392308443784714
step: 160, loss: 0.07059699296951294
step: 170, loss: 0.02354983612895012
step: 180, loss: 0.05355813726782799
step: 190, loss: 0.10480744391679764
step: 200, loss: 0.15159663558006287
step: 210, loss: 0.0752185806632042
/usr/local/lib/python3.7/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=0.690909090909091, f1=0.7266055045871559,
best_f1=0.7446808510638298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06825831532478333
step: 10, loss: 0.002876454498618841

```

```
step: 20, loss: 0.17318591475486755
step: 30, loss: 0.03817647695541382
step: 40, loss: 0.029429908841848373
step: 50, loss: 0.035529181361198425
step: 60, loss: 0.03609108179807663
step: 70, loss: 0.06532584130764008
step: 80, loss: 0.035032376646995544
step: 90, loss: 0.0869114026427269
step: 100, loss: 0.04484553262591362
step: 110, loss: 0.12308502197265625
step: 120, loss: 0.17460396885871887
step: 130, loss: 0.02291218750178814
step: 140, loss: 0.030653387308120728
step: 150, loss: 0.02577594481408596
step: 160, loss: 0.21282345056533813
step: 170, loss: 0.18789120018482208
step: 180, loss: 0.02353283204138279
step: 190, loss: 0.1723771095275879
step: 200, loss: 0.030076023191213608
step: 210, loss: 0.0906769186258316
/usr/local/lib/python3.7/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=0.6974169741697417, f1=0.716636197440585,
best_f1=0.7446808510638298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18293331563472748
step: 10, loss: 0.014271813444793224
step: 20, loss: 0.06395319849252701
step: 30, loss: 0.030017007142305374
step: 40, loss: 0.005456532351672649
step: 50, loss: 0.000803700415417552
step: 60, loss: 0.024897435680031776
step: 70, loss: 0.026892224326729774
step: 80, loss: 0.10997192561626434
step: 90, loss: 0.03745054826140404
step: 100, loss: 0.3165733218193054
step: 110, loss: 0.01247031893581152
step: 120, loss: 0.048098355531692505
step: 130, loss: 0.007319482043385506
step: 140, loss: 0.05682311952114105
step: 150, loss: 0.12792333960533142
step: 160, loss: 0.12259411066770554
step: 170, loss: 0.17423534393310547
step: 180, loss: 0.06138370931148529
step: 190, loss: 0.0119937714189291
```

```

step: 200, loss: 0.0658847987651825
step: 210, loss: 0.06844909489154816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7123809523809524, f1=0.7392996108949417,
best_f1=0.7392996108949417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09133310616016388
step: 10, loss: 0.13124093413352966
step: 20, loss: 0.07631618529558182
step: 30, loss: 0.03688371554017067
step: 40, loss: 0.013222538866102695
step: 50, loss: 0.017929092049598694
step: 60, loss: 0.005436718463897705
step: 70, loss: 0.1272323578596115
step: 80, loss: 0.00414709048345685
step: 90, loss: 0.020325608551502228
step: 100, loss: 0.051013655960559845
step: 110, loss: 0.06493698060512543
step: 120, loss: 0.16371284425258636
step: 130, loss: 0.0718674287199974
step: 140, loss: 0.06795153021812439
step: 150, loss: 0.00733413128182292
step: 160, loss: 0.20675760507583618
step: 170, loss: 0.04888836294412613
step: 180, loss: 0.036152858287096024
step: 190, loss: 0.015809321776032448
step: 200, loss: 0.008923957124352455
step: 210, loss: 0.08013395965099335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7164750957854407, f1=0.745631067961165,
best_f1=0.745631067961165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.012565779499709606
step: 10, loss: 0.009649870917201042
step: 20, loss: 0.11365509033203125
step: 30, loss: 0.024776864796876907
step: 40, loss: 0.10756688565015793
step: 50, loss: 0.05383804813027382
step: 60, loss: 0.041693929582834244
step: 70, loss: 0.01866387575864792
step: 80, loss: 0.20191101729869843
step: 90, loss: 0.017968395724892616
step: 100, loss: 0.06692834198474884
step: 110, loss: 0.07020303606987
step: 120, loss: 0.06662079691886902
step: 130, loss: 0.06493821740150452
step: 140, loss: 0.0019911997951567173
step: 150, loss: 0.009936882182955742
step: 160, loss: 0.007995347492396832
step: 170, loss: 0.0317440927028656
step: 180, loss: 0.055399078875780106
step: 190, loss: 0.05361533537507057
step: 200, loss: 0.033883340656757355
step: 210, loss: 0.028101343661546707
/usr/local/lib/python3.7/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=0.7054108216432866, f1=0.7221095334685598,
best_f1=0.745631067961165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04651635140180588
step: 10, loss: 0.06068871170282364
step: 20, loss: 0.049980662763118744
step: 30, loss: 0.015892228111624718
step: 40, loss: 0.012652448378503323
step: 50, loss: 0.0018995825666934252
step: 60, loss: 0.02820245735347271
step: 70, loss: 0.03065394051373005
step: 80, loss: 0.04981313645839691
step: 90, loss: 0.09340070188045502
step: 100, loss: 0.01773971877992153
step: 110, loss: 0.13070231676101685
step: 120, loss: 0.023447232320904732
step: 130, loss: 0.005164121277630329
step: 140, loss: 0.0812186673283577
step: 150, loss: 0.11400918662548065
step: 160, loss: 0.010511904954910278
step: 170, loss: 0.02708345651626587

```

```

step: 180, loss: 0.0015349400928243995
step: 190, loss: 0.05045219883322716
step: 200, loss: 0.0155861871317029
step: 210, loss: 0.005013987421989441
/usr/local/lib/python3.7/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=0.7100371747211897, f1=0.7245283018867925,
best_f1=0.745631067961165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03649599105119705
step: 10, loss: 0.04275965318083763
step: 20, loss: 0.10523581504821777
step: 30, loss: 0.008135592564940453
step: 40, loss: 0.0006640172796323895
step: 50, loss: 0.006763802375644445
step: 60, loss: 0.005265276413410902
step: 70, loss: 0.007730788551270962
step: 80, loss: 0.2755535840988159
step: 90, loss: 0.020518114790320396
step: 100, loss: 0.02894454076886177
step: 110, loss: 0.03293319791555405
step: 120, loss: 0.0004227012977935374
step: 130, loss: 0.01399753987789154
step: 140, loss: 0.2555240988731384
step: 150, loss: 0.0011451146565377712
step: 160, loss: 0.011441754177212715
step: 170, loss: 0.05235694348812103
step: 180, loss: 0.007669584359973669
step: 190, loss: 0.03395616263151169
step: 200, loss: 0.011760162189602852
step: 210, loss: 0.0015655708266422153
/usr/local/lib/python3.7/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=0.7140115163147792, f1=0.7366336633663367,
best_f1=0.745631067961165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03570285066962242
step: 10, loss: 0.11203666031360626
step: 20, loss: 0.0007920144707895815
step: 30, loss: 0.0006845143507234752
step: 40, loss: 0.011398255825042725
step: 50, loss: 0.05038543418049812

```

```

step: 60, loss: 0.009993623942136765
step: 70, loss: 0.003012656467035413
step: 80, loss: 0.28163114190101624
step: 90, loss: 0.014557705260813236
step: 100, loss: 0.002682457445189357
step: 110, loss: 0.022496959194540977
step: 120, loss: 0.0982607826590538
step: 130, loss: 0.000409311440307647
step: 140, loss: 0.02663266658782959
step: 150, loss: 0.0019098938209936023
step: 160, loss: 0.08586463332176208
step: 170, loss: 0.028225954622030258
step: 180, loss: 0.00788089632987976
step: 190, loss: 0.03168454393744469
step: 200, loss: 0.040766384452581406
step: 210, loss: 0.125019833445549
/usr/local/lib/python3.7/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=0.7106796116504855, f1=0.74, best_f1=0.745631067961165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007872692309319973
step: 10, loss: 0.06589113175868988
step: 20, loss: 0.000817907159216702
step: 30, loss: 0.002277834340929985
step: 40, loss: 0.015560596249997616
step: 50, loss: 0.003417645115405321
step: 60, loss: 0.037744127213954926
step: 70, loss: 0.009369530715048313
step: 80, loss: 0.013627487234771252
step: 90, loss: 0.0012331954203546047
step: 100, loss: 0.00945243425667286
step: 110, loss: 0.0028448086231946945
step: 120, loss: 0.0029070349410176277
step: 130, loss: 0.03222817927598953
step: 140, loss: 0.0015805402072146535
step: 150, loss: 0.054124992340803146
step: 160, loss: 0.005808528512716293
step: 170, loss: 0.08384525030851364
step: 180, loss: 0.004855472594499588
step: 190, loss: 0.04539636895060539
step: 200, loss: 0.07115646451711655
step: 210, loss: 0.0009489592630416155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7215686274509804, f1=0.7410358565737052,
best_f1=0.7410358565737052
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015503731556236744
step: 10, loss: 0.0014670321252197027
step: 20, loss: 0.05654929205775261
step: 30, loss: 0.012103430926799774
step: 40, loss: 0.010682076215744019
step: 50, loss: 0.01020123902708292
step: 60, loss: 0.030027318745851517
step: 70, loss: 0.023287974298000336
step: 80, loss: 0.07318129390478134
step: 90, loss: 0.003703808644786477
step: 100, loss: 0.01403928454965353
step: 110, loss: 0.042660173028707504
step: 120, loss: 0.010555002838373184
step: 130, loss: 0.012004492804408073
step: 140, loss: 0.0485261008143425
step: 150, loss: 0.11101844906806946
step: 160, loss: 0.01838483288884163
step: 170, loss: 0.02298332005739212
step: 180, loss: 0.0010316550033167005
step: 190, loss: 0.003486644010990858
step: 200, loss: 0.0003899727307725698
step: 210, loss: 0.01741557940840721
/usr/local/lib/python3.7/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=0.7185628742514971, f1=0.7280163599182004,
best_f1=0.7410358565737052
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, 240.62it/s]
load_f1 = 0.7414448669201521
real_f1 = 0.7324478178368121
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, 146.00it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.44920971989631653
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.38577353954315186
step: 20, loss: 0.2541659474372864
step: 30, loss: 0.3948453366756439
step: 40, loss: 0.2686896324157715
step: 50, loss: 0.32565316557884216
step: 60, loss: 0.4354420304298401
step: 70, loss: 0.4210752546787262
step: 80, loss: 0.22458374500274658
step: 90, loss: 0.2821667492389679
step: 100, loss: 0.40916672348976135
step: 110, loss: 0.23106339573860168
step: 120, loss: 0.2893627882003784
step: 130, loss: 0.30864736437797546
step: 140, loss: 0.11549234390258789
step: 150, loss: 0.21966063976287842
step: 160, loss: 0.15514475107192993
step: 170, loss: 0.48845019936561584
step: 180, loss: 0.17909395694732666
step: 190, loss: 0.12987488508224487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5706940874035991, f1=0.5927710843373495,
best_f1=0.5927710843373495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24051865935325623
step: 10, loss: 0.12268547713756561
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.7729182243347168
step: 30, loss: 0.16271495819091797
step: 40, loss: 0.36159658432006836
step: 50, loss: 0.26702871918678284
step: 60, loss: 0.18767321109771729
step: 70, loss: 0.23362024128437042
step: 80, loss: 0.1542285680770874
```

```

step: 90, loss: 0.13985829055309296
step: 100, loss: 0.0788247138261795
step: 110, loss: 0.1531437337398529
step: 120, loss: 0.14124295115470886
step: 130, loss: 0.15384627878665924
step: 140, loss: 0.14948667585849762
step: 150, loss: 0.2039574384689331
step: 160, loss: 0.19119103252887726
step: 170, loss: 0.11319841444449234
step: 180, loss: 0.03631923720240593
step: 190, loss: 0.2035648673772812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7253886010362695, f1=0.7539267015706806,
best_f1=0.7539267015706806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24662156403064728
step: 10, loss: 0.1449965238571167
step: 20, loss: 0.06032266840338707
step: 30, loss: 0.02477399818599224
step: 40, loss: 0.09368517994880676
step: 50, loss: 0.11667543649673462
step: 60, loss: 0.06227678805589676
step: 70, loss: 0.09083130955696106
step: 80, loss: 0.09013904631137848
step: 90, loss: 0.16098663210868835
step: 100, loss: 0.0412106066942215
step: 110, loss: 0.5297395586967468
step: 120, loss: 0.18742632865905762
step: 130, loss: 0.06248002499341965
step: 140, loss: 0.03318000212311745
step: 150, loss: 0.15058597922325134
step: 160, loss: 0.080033078789711
step: 170, loss: 0.17863328754901886
step: 180, loss: 0.16001375019550323
step: 190, loss: 0.01712222211062908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8379888268156425, f1=0.8022284122562673,
best_f1=0.8022284122562673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006428753025829792
step: 10, loss: 0.09140337258577347
step: 20, loss: 0.09176938980817795
step: 30, loss: 0.024951491504907608
step: 40, loss: 0.09457338601350784
step: 50, loss: 0.025967461988329887
step: 60, loss: 0.012177631258964539
step: 70, loss: 0.030034605413675308
step: 80, loss: 0.037407368421554565
step: 90, loss: 0.167541041970253
step: 100, loss: 0.19489368796348572
step: 110, loss: 0.3397144675254822
step: 120, loss: 0.03508437052369118
step: 130, loss: 0.06189781799912453
step: 140, loss: 0.30744513869285583
step: 150, loss: 0.11907359212636948
step: 160, loss: 0.041722554713487625
step: 170, loss: 0.04445226490497589
step: 180, loss: 0.03599371388554573
step: 190, loss: 0.014144529588520527
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.842391304347826, f1=0.8401084010840109,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043352678418159485
step: 10, loss: 0.14254902303218842
step: 20, loss: 0.013710584491491318
step: 30, loss: 0.04564506933093071
step: 40, loss: 0.0869104415178299
step: 50, loss: 0.022925181314349174
step: 60, loss: 0.06450178474187851
step: 70, loss: 0.09428904205560684
step: 80, loss: 0.022929061204195023

```

```

step: 90, loss: 0.05026257038116455
step: 100, loss: 0.03378058224916458
step: 110, loss: 0.243287593126297
step: 120, loss: 0.06670048832893372
step: 130, loss: 0.23443962633609772
step: 140, loss: 0.025203775614500046
step: 150, loss: 0.07031184434890747
step: 160, loss: 0.007939095608890057
step: 170, loss: 0.06306136399507523
step: 180, loss: 0.013992352411150932
step: 190, loss: 0.03867128863930702
/usr/local/lib/python3.7/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=0.8368421052631578, f1=0.8481675392670156,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2473514974117279
step: 10, loss: 0.01745537854731083
step: 20, loss: 0.006764921359717846
step: 30, loss: 0.028796803206205368
step: 40, loss: 0.036527276039123535
step: 50, loss: 0.03002282977104187
step: 60, loss: 0.073723204433918
step: 70, loss: 0.2065179944038391
step: 80, loss: 0.03542295843362808
step: 90, loss: 0.06076262891292572
step: 100, loss: 0.10988274961709976
step: 110, loss: 0.013314256444573402
step: 120, loss: 0.01774836890399456
step: 130, loss: 0.11021210253238678
step: 140, loss: 0.00620344327762723
step: 150, loss: 0.00616824347525835
step: 160, loss: 0.11899663507938385
step: 170, loss: 0.13925901055335999
step: 180, loss: 0.049980565905570984
step: 190, loss: 0.016821658238768578
/usr/local/lib/python3.7/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=0.8264462809917356, f1=0.827027027027027,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028997836634516716

```

```

step: 10, loss: 0.0019561557564884424
step: 20, loss: 0.0018355833599343896
step: 30, loss: 0.028838476166129112
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 40, loss: 0.0180191732943058
step: 50, loss: 0.013851854018867016
step: 60, loss: 0.09398264437913895
step: 70, loss: 0.0015264945104718208
step: 80, loss: 0.01320702489465475
step: 90, loss: 0.029629549011588097
step: 100, loss: 0.16088983416557312
step: 110, loss: 0.07549474388360977
step: 120, loss: 0.025261476635932922
step: 130, loss: 0.00765474047511816
step: 140, loss: 0.02038533426821232
step: 150, loss: 0.023759255185723305
step: 160, loss: 0.03492512181401253
step: 170, loss: 0.05228137969970703
step: 180, loss: 0.003957520239055157
step: 190, loss: 0.04667320102453232
/usr/local/lib/python3.7/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=0.8302872062663187, f1=0.8093994778067886,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.049709923565387726
step: 10, loss: 0.013428712263703346
step: 20, loss: 0.015092620626091957
step: 30, loss: 0.024478629231452942
step: 40, loss: 0.03011457249522209
step: 50, loss: 0.043831080198287964
step: 60, loss: 0.03152650594711304
step: 70, loss: 0.004420614335685968
step: 80, loss: 0.06874046474695206
step: 90, loss: 0.04312766343355179
step: 100, loss: 0.02271527238190174
step: 110, loss: 0.05038021132349968
step: 120, loss: 0.0023091472685337067
step: 130, loss: 0.00647427374497056
step: 140, loss: 0.0027086858171969652
step: 150, loss: 0.040570519864559174
step: 160, loss: 0.004296866245567799
step: 170, loss: 0.0006759088719263673
step: 180, loss: 0.005053908564150333
step: 190, loss: 0.0019690440967679024

```

```

/usr/local/lib/python3.7/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=0.8279569892473119, f1=0.8369565217391304,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005378908826969564
step: 10, loss: 0.001605480327270925
step: 20, loss: 0.0036474401131272316
step: 30, loss: 0.04237920790910721
step: 40, loss: 0.02015901356935501
step: 50, loss: 0.01454502996057272
step: 60, loss: 0.005945621524006128
step: 70, loss: 0.0044528041034936905
step: 80, loss: 0.011900225654244423
step: 90, loss: 0.007353195454925299
step: 100, loss: 0.008447620086371899
step: 110, loss: 0.0018830876797437668
step: 120, loss: 0.15608695149421692
step: 130, loss: 0.010291964747011662
step: 140, loss: 0.12124060094356537
step: 150, loss: 0.00424768915399909
step: 160, loss: 0.002460319548845291
step: 170, loss: 0.02419554442167282
step: 180, loss: 0.013511420227587223
step: 190, loss: 0.0010594545165076852
/usr/local/lib/python3.7/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=0.8398950131233596, f1=0.8373333333333334,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005289938999339938
step: 10, loss: 0.0011299633188173175
step: 20, loss: 0.0024074865505099297
step: 30, loss: 0.03924667090177536
step: 40, loss: 0.004672623705118895
step: 50, loss: 0.0018890637438744307
step: 60, loss: 0.001925283344462514
step: 70, loss: 0.0020670879166573286
step: 80, loss: 0.005470046307891607
step: 90, loss: 0.0015531207900494337
step: 100, loss: 0.04080716520547867
step: 110, loss: 0.003161142347380519

```

```

step: 120, loss: 0.009164527989923954
step: 130, loss: 0.04091303050518036
step: 140, loss: 0.0014994984958320856
step: 150, loss: 0.010864818468689919
step: 160, loss: 0.0036029841285198927
step: 170, loss: 0.0011026873253285885
step: 180, loss: 0.10559652000665665
step: 190, loss: 0.0003184294910170138
/usr/local/lib/python3.7/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=0.8410256410256409, f1=0.8527131782945736,
best_f1=0.8401084010840109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01080489344894886
step: 10, loss: 0.015403305180370808
step: 20, loss: 0.0018705392722040415
step: 30, loss: 0.03662360459566116
step: 40, loss: 0.0006532236002385616
step: 50, loss: 0.15355482697486877
step: 60, loss: 0.014322792179882526
step: 70, loss: 0.0038049265276640654
step: 80, loss: 0.1236824318766594
step: 90, loss: 0.0051774219609797
step: 100, loss: 0.02450445666909218
step: 110, loss: 0.21193064749240875
step: 120, loss: 0.0014222182799130678
step: 130, loss: 0.006029897835105658
step: 140, loss: 0.0020424926187843084
step: 150, loss: 0.0010682385182008147
step: 160, loss: 0.0028737233951687813
step: 170, loss: 0.0022387877106666565
step: 180, loss: 0.006643849425017834
step: 190, loss: 0.0088027473539114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8571428571428572, f1=0.8505154639175257,
best_f1=0.8505154639175257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.02614523656666279
step: 10, loss: 0.010452200658619404
step: 20, loss: 0.00280755921266973
step: 30, loss: 0.06673870235681534
step: 40, loss: 0.0030972028616815805
step: 50, loss: 0.0010343979811295867
step: 60, loss: 0.00508946692571044
step: 70, loss: 0.04066735878586769
step: 80, loss: 0.002041239757090807
step: 90, loss: 0.001243499224074185
step: 100, loss: 0.010865709744393826
step: 110, loss: 0.046097781509160995
step: 120, loss: 0.007653036154806614
step: 130, loss: 0.007327501196414232
step: 140, loss: 0.0028352837543934584
step: 150, loss: 0.012129341252148151
step: 160, loss: 0.00297686574049294
step: 170, loss: 0.030730340629816055
step: 180, loss: 0.000764673575758934
step: 190, loss: 0.002984563587233424
/usr/local/lib/python3.7/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=0.851063829787234, f1=0.8337730870712402,
best_f1=0.8505154639175257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020898113027215004
step: 10, loss: 0.01518232375383377
step: 20, loss: 0.0010392561089247465
step: 30, loss: 0.08047781884670258
step: 40, loss: 0.0028633479960262775
step: 50, loss: 0.004252285696566105
step: 60, loss: 0.03997024893760681
step: 70, loss: 0.024315237998962402
step: 80, loss: 0.0015609932597726583
step: 90, loss: 0.0070021203719079494
step: 100, loss: 0.004404302220791578
step: 110, loss: 0.0016809345688670874
step: 120, loss: 0.00042165303602814674
step: 130, loss: 0.003062720876187086
step: 140, loss: 0.0033400405663996935
step: 150, loss: 0.0017712890403345227
step: 160, loss: 0.0027320918161422014
step: 170, loss: 0.004615072160959244
step: 180, loss: 0.0010389366652816534
step: 190, loss: 0.0950704962015152

```

```

/usr/local/lib/python3.7/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=0.8463611859838275, f1=0.8455284552845528,
best_f1=0.8505154639175257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008547661826014519
step: 10, loss: 0.02912590652704239
step: 20, loss: 0.0075081163085997105
step: 30, loss: 0.0014148958725854754
step: 40, loss: 0.00028610462322831154
step: 50, loss: 0.0009399748523719609
step: 60, loss: 0.001460377243347466
step: 70, loss: 0.001852123998105526
step: 80, loss: 0.0002187543432282672
step: 90, loss: 0.00019234835053794086
step: 100, loss: 0.0005172837991267443
step: 110, loss: 0.00017445812409278005
step: 120, loss: 0.03148354962468147
step: 130, loss: 0.00021600090258289129
step: 140, loss: 0.0014766649110242724
step: 150, loss: 0.002696473151445389
step: 160, loss: 0.0036478876136243343
step: 170, loss: 0.0022623897530138493
step: 180, loss: 0.0014660459710285068
step: 190, loss: 0.0009284861152991652
/usr/local/lib/python3.7/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=0.8463611859838275, f1=0.8478260869565217,
best_f1=0.8505154639175257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000770131649915129
step: 10, loss: 0.011297749355435371
step: 20, loss: 0.01685182936489582
step: 30, loss: 0.001341083669103682
step: 40, loss: 0.0038024557288736105
step: 50, loss: 0.0003129554388578981
step: 60, loss: 0.0006753710913471878
step: 70, loss: 0.010021514259278774
step: 80, loss: 0.01413747202605009
step: 90, loss: 0.0033179912716150284
step: 100, loss: 0.0007491864380426705
step: 110, loss: 0.005744622554630041

```

```

step: 120, loss: 0.0010575815103948116
step: 130, loss: 0.007687482517212629
step: 140, loss: 0.10010125488042831
step: 150, loss: 0.0003089855599682778
step: 160, loss: 0.0003693479229696095
step: 170, loss: 0.006200138013809919
step: 180, loss: 0.0002723791403695941
step: 190, loss: 0.0027417331002652645
/usr/local/lib/python3.7/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=0.8478260869565217,
best_f1=0.8505154639175257
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:13, 154.07it/s]
load_f1 = 0.7772277227227273
real_f1 = 0.7547169811320754
733it [00:00, 3354.19it/s]/usr/local/lib/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, 130.14it/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/
```

```
Downloading: 100% 481/481 [00:00<00:00, 426kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 23.7MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 23.8MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 69.8MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
```

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

/usr/local/lib/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: 0.4809724986553192

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

step: 10, loss: 0.44147467613220215

step: 20, loss: 0.31220340728759766

step: 30, loss: 0.3811098039150238

step: 40, loss: 0.537312924861908

step: 50, loss: 0.3327130079269409

step: 60, loss: 0.5393364429473877

step: 70, loss: 0.3226163983345032

step: 80, loss: 0.26091837882995605

step: 90, loss: 0.22560130059719086

step: 100, loss: 0.16713497042655945

step: 110, loss: 0.39238062500953674

```

step: 120, loss: 0.32941558957099915
step: 130, loss: 0.31087347865104675
step: 140, loss: 0.3958001136779785
step: 150, loss: 0.30497869849205017
step: 160, loss: 0.3678394556045532
step: 170, loss: 0.2035014033317566
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6804597701149426, f1=0.687089715536105,
best_f1=0.687089715536105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1881149560213089
step: 10, loss: 0.17914935946464539
step: 20, loss: 0.13098561763763428
step: 30, loss: 0.23276148736476898
step: 40, loss: 0.09349776059389114
step: 50, loss: 0.17131231725215912
step: 60, loss: 0.028890758752822876
step: 70, loss: 0.1430109292268753
step: 80, loss: 0.16013282537460327
step: 90, loss: 0.040234602987766266
step: 100, loss: 0.07481750100851059
step: 110, loss: 0.1020185649394989
step: 120, loss: 0.017044037580490112
step: 130, loss: 0.2681496739387512
step: 140, loss: 0.21886761486530304
step: 150, loss: 0.15398883819580078
step: 160, loss: 0.17036235332489014
step: 170, loss: 0.1371927410364151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8353808353808354, f1=0.8591549295774649,
best_f1=0.8591549295774649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11627976596355438
step: 10, loss: 0.025609511882066727
step: 20, loss: 0.09149618446826935
step: 30, loss: 0.0048131998628377914
step: 40, loss: 0.15448321402072906
step: 50, loss: 0.2900669574737549
step: 60, loss: 0.06887157261371613
step: 70, loss: 0.23287653923034668
step: 80, loss: 0.07346007227897644
step: 90, loss: 0.19986926019191742
step: 100, loss: 0.008320377208292484
step: 110, loss: 0.05481633543968201
step: 120, loss: 0.04651900380849838
step: 130, loss: 0.03980530425906181
step: 140, loss: 0.13582183420658112
step: 150, loss: 0.03509402275085449
step: 160, loss: 0.02098945714533329
step: 170, loss: 0.01153674814850092
/usr/local/lib/python3.7/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=0.82903981264637, f1=0.8701594533029613,
best_f1=0.8591549295774649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044867753982543945
step: 10, loss: 0.0033200220204889774
step: 20, loss: 0.0035822719801217318
step: 30, loss: 0.09790297597646713
step: 40, loss: 0.07671818137168884
step: 50, loss: 0.03220256790518761
step: 60, loss: 0.14573664963245392
step: 70, loss: 0.0015579629689455032
step: 80, loss: 0.18255925178527832
step: 90, loss: 0.17934487760066986
step: 100, loss: 0.11077689379453659
step: 110, loss: 0.204390749335289
step: 120, loss: 0.14143899083137512
step: 130, loss: 0.09844208508729935
step: 140, loss: 0.048382263630628586
step: 150, loss: 0.17351652681827545
step: 160, loss: 0.025377480313181877
step: 170, loss: 0.05434408411383629
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8731707317073171, f1=0.8831775700934579,
best_f1=0.8831775700934579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022691993042826653
step: 10, loss: 0.017646340653300285
step: 20, loss: 0.03058508411049843
step: 30, loss: 0.006005147472023964
step: 40, loss: 0.02814677357673645
step: 50, loss: 0.010810203850269318
step: 60, loss: 0.06826498359441757
step: 70, loss: 0.07011353224515915
step: 80, loss: 0.007121887058019638
step: 90, loss: 0.056381527334451675
step: 100, loss: 0.06322891265153885
step: 110, loss: 0.03764912113547325
step: 120, loss: 0.09157723188400269
step: 130, loss: 0.0017244888003915548
step: 140, loss: 0.030954955145716667
step: 150, loss: 0.022055653855204582
step: 160, loss: 0.010301872156560421
step: 170, loss: 0.03471623733639717
/usr/local/lib/python3.7/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=0.860759493670886, f1=0.8905472636815919,
best_f1=0.8831775700934579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01242861058562994
step: 10, loss: 0.007805068511515856
step: 20, loss: 0.07271692901849747
step: 30, loss: 0.011421873234212399
step: 40, loss: 0.003366826567798853
step: 50, loss: 0.010076886974275112
step: 60, loss: 0.01211739145219326
step: 70, loss: 0.034402236342430115
step: 80, loss: 0.023351915180683136
step: 90, loss: 0.057396434247493744
step: 100, loss: 0.01217649132013321
step: 110, loss: 0.06257206946611404
step: 120, loss: 0.07897453010082245

```

```

step: 130, loss: 0.037043239921331406
step: 140, loss: 0.043861035257577896
step: 150, loss: 0.006932167802006006
step: 160, loss: 0.038398634642362595
step: 170, loss: 0.001929199555888772
/usr/local/lib/python3.7/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=0.8702290076335877, f1=0.8960396039603961,
best_f1=0.8831775700934579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03566865250468254
step: 10, loss: 0.04223523288965225
step: 20, loss: 0.017393801361322403
step: 30, loss: 0.003428169758990407
step: 40, loss: 0.00026361801428720355
step: 50, loss: 0.0002765964309219271
step: 60, loss: 0.0013679546536877751
step: 70, loss: 0.045216288417577744
step: 80, loss: 0.07799319177865982
step: 90, loss: 0.03645138069987297
step: 100, loss: 0.0008547697216272354
step: 110, loss: 0.07583072781562805
step: 120, loss: 0.005798752419650555
step: 130, loss: 0.012756251730024815
step: 140, loss: 0.004316739737987518
step: 150, loss: 0.15578824281692505
step: 160, loss: 0.008635543286800385
step: 170, loss: 0.0680222138762474
/usr/local/lib/python3.7/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=0.8563829787234041, f1=0.8804071246819339,
best_f1=0.8831775700934579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04171279817819595
step: 10, loss: 0.015746043995022774
step: 20, loss: 0.0035176172386854887
step: 30, loss: 0.027138635516166687
step: 40, loss: 0.0007027764804661274
step: 50, loss: 0.0008936601807363331
step: 60, loss: 0.005732121877372265
step: 70, loss: 0.020698389038443565
step: 80, loss: 0.049911316484212875

```

```

step: 90, loss: 0.14145098626613617
step: 100, loss: 0.001181631232611835
step: 110, loss: 0.11604303121566772
step: 120, loss: 0.004249549005180597
step: 130, loss: 0.0019416494760662317
step: 140, loss: 0.03671306371688843
step: 150, loss: 0.0008102747960947454
step: 160, loss: 0.015652084723114967
step: 170, loss: 0.14974698424339294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8768472906403941, f1=0.892018779342723,
best_f1=0.892018779342723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023796873167157173
step: 10, loss: 0.005915238056331873
step: 20, loss: 0.0013414763379842043
step: 30, loss: 0.0007046476239338517
step: 40, loss: 0.0034343490842729807
step: 50, loss: 0.0004910871502943337
step: 60, loss: 0.001356205204501748
step: 70, loss: 0.00020635331748053432
step: 80, loss: 0.0010811473475769162
step: 90, loss: 0.009733827784657478
step: 100, loss: 0.0476977676153183
step: 110, loss: 0.010521532036364079
step: 120, loss: 0.001691198442131281
step: 130, loss: 0.003287010593339801
step: 140, loss: 0.0031903134658932686
step: 150, loss: 0.1157074123620987
step: 160, loss: 0.002545570023357868
step: 170, loss: 0.05260062590241432
/usr/local/lib/python3.7/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=0.8636363636363636, f1=0.8997555012224939,
best_f1=0.892018779342723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001802691724151373

```

```

step: 10, loss: 0.011048860847949982
step: 20, loss: 0.03556356579065323
step: 30, loss: 0.07223211973905563
step: 40, loss: 0.0010270107304677367
step: 50, loss: 0.00435035303235054
step: 60, loss: 0.0062004271894693375
step: 70, loss: 0.01204984076321125
step: 80, loss: 0.021449461579322815
step: 90, loss: 0.000264892092673108
step: 100, loss: 0.019375041127204895
step: 110, loss: 0.06787115335464478
step: 120, loss: 0.00039319245843216777
step: 130, loss: 0.0035851639695465565
step: 140, loss: 0.002168821869418025
step: 150, loss: 0.02272217720746994
step: 160, loss: 0.00014812541485298425
step: 170, loss: 0.19665582478046417
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8778054862842892, f1=0.8957345971563981,
best_f1=0.8957345971563981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014984220266342163
step: 10, loss: 0.028608310967683792
step: 20, loss: 0.0009031532681547105
step: 30, loss: 0.040087077766656876
step: 40, loss: 0.00042913469951599836
step: 50, loss: 0.0032099017407745123
step: 60, loss: 0.014783656224608421
step: 70, loss: 0.0053672268986701965
step: 80, loss: 0.051805175840854645
step: 90, loss: 0.005315111950039864
step: 100, loss: 0.04432756453752518
step: 110, loss: 0.011412850581109524
step: 120, loss: 0.0059811705723404884
step: 130, loss: 0.023746374994516373
step: 140, loss: 0.0005707753589376807
step: 150, loss: 0.010577701963484287
step: 160, loss: 0.0012573805870488286
step: 170, loss: 0.027490366250276566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8872180451127819, f1=0.905109489051095,
best_f1=0.905109489051095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013132572639733553
step: 10, loss: 0.001448746770620346
step: 20, loss: 0.0008969012997113168
step: 30, loss: 0.04080379381775856
step: 40, loss: 0.03261040523648262
step: 50, loss: 0.0024365861900150776
step: 60, loss: 0.09457437694072723
step: 70, loss: 0.002640407532453537
step: 80, loss: 0.00031242711702361703
step: 90, loss: 0.0010082360822707415
step: 100, loss: 0.0002632547984831035
step: 110, loss: 0.03855578973889351
step: 120, loss: 0.02848966233432293
step: 130, loss: 0.048935968428850174
step: 140, loss: 0.002585401525720954
step: 150, loss: 0.0022749758791178465
step: 160, loss: 0.0014956592349335551
step: 170, loss: 0.0007972203311510384
/usr/local/lib/python3.7/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=0.8717948717948718, f1=0.9077669902912622,
best_f1=0.905109489051095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002557237632572651
step: 10, loss: 0.0010636409278959036
step: 20, loss: 0.0005390621372498572
step: 30, loss: 0.04069957882165909
step: 40, loss: 0.006557210348546505
step: 50, loss: 0.00935774203389883
step: 60, loss: 0.0007387559744529426
step: 70, loss: 0.06110541522502899
step: 80, loss: 0.0009811759227886796
step: 90, loss: 0.002579373773187399
step: 100, loss: 0.02622305601835251

```

```

step: 110, loss: 0.0011271871626377106
step: 120, loss: 0.013820815831422806
step: 130, loss: 0.0015667150728404522
step: 140, loss: 0.017255987972021103
step: 150, loss: 0.0004304300819057971
step: 160, loss: 0.0003447247145231813
step: 170, loss: 0.0022924963850528
/usr/local/lib/python3.7/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=0.8727272727272727, f1=0.9014778325123154,
best_f1=0.905109489051095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09139115363359451
step: 10, loss: 0.05623330548405647
step: 20, loss: 0.0012709434377029538
step: 30, loss: 0.0007139961235225201
step: 40, loss: 0.006341839209198952
step: 50, loss: 0.06435438990592957
step: 60, loss: 0.010972581803798676
step: 70, loss: 0.008614457212388515
step: 80, loss: 0.002249770099297166
step: 90, loss: 0.00039586081402376294
step: 100, loss: 0.00030760595109313726
step: 110, loss: 0.020723186433315277
step: 120, loss: 0.011345887556672096
step: 130, loss: 0.029689867049455643
step: 140, loss: 0.020616142079234123
step: 150, loss: 0.02075052633881569
step: 160, loss: 0.0009636614122428
step: 170, loss: 0.01942077837884426
/usr/local/lib/python3.7/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=0.884318766066838, f1=0.9073170731707318,
best_f1=0.905109489051095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003696685307659209
step: 10, loss: 0.0004463512450456619
step: 20, loss: 0.0013695688685402274
step: 30, loss: 0.0011701115872710943
step: 40, loss: 0.00034739464172162116
step: 50, loss: 0.0029026977717876434
step: 60, loss: 0.0013399816816672683

```

```

step: 70, loss: 0.0804348960518837
step: 80, loss: 0.0012473969254642725
step: 90, loss: 0.0013105073012411594
step: 100, loss: 0.039503514766693115
step: 110, loss: 0.00025617689243517816
step: 120, loss: 0.0003359418478794396
step: 130, loss: 0.0003461361047811806
step: 140, loss: 0.024643585085868835
step: 150, loss: 0.0002438496012473479
step: 160, loss: 0.0006750928005203605
step: 170, loss: 0.07294911891222
/usr/local/lib/python3.7/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=0.881443298969072, f1=0.9095354523227384,
best_f1=0.905109489051095
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:05, 322.07it/s]
load_f1 = 0.4768041237113402

```

```

real_f1 = 0.4258210645526614
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 196.31it/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 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.")
step: 0, loss: 0.5604046583175659
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: 0.4835669696331024
step: 20, loss: 0.43725907802581787
step: 30, loss: 0.28600677847862244
step: 40, loss: 0.20382574200630188
step: 50, loss: 0.25866004824638367
step: 60, loss: 0.08690299093723297
step: 70, loss: 0.2298106551170349
step: 80, loss: 0.04565480723977089
step: 90, loss: 0.2244413048028946
step: 100, loss: 0.1039016917347908
step: 110, loss: 0.12750710546970367
step: 120, loss: 0.081343874335289
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.010834112763404846
step: 140, loss: 0.03905317187309265
step: 150, loss: 0.1100766509771347
step: 160, loss: 0.02665596641600132
step: 170, loss: 0.09681466221809387
step: 180, loss: 0.2544960677623749
step: 190, loss: 0.048075269907712936
step: 200, loss: 0.05451397970318794

```

```

step: 210, loss: 0.04593772441148758
step: 220, loss: 0.06628558784723282
step: 230, loss: 0.17279799282550812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9496717724288839, f1=0.9612403100775194,
best_f1=0.9612403100775194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02889905869960785
step: 10, loss: 0.14939965307712555
step: 20, loss: 0.024579893797636032
step: 30, loss: 0.12411927431821823
step: 40, loss: 0.061315592378377914
step: 50, loss: 0.006646266672760248
step: 60, loss: 0.004244385752826929
step: 70, loss: 0.08631866425275803
step: 80, loss: 0.007967432960867882
step: 90, loss: 0.005224103108048439
step: 100, loss: 0.12364183366298676
step: 110, loss: 0.08888070285320282
step: 120, loss: 0.007289126981049776
step: 130, loss: 0.018261563032865524
step: 140, loss: 0.004696684889495373
step: 150, loss: 0.21516810357570648
step: 160, loss: 0.01658105105161667
step: 170, loss: 0.003916415385901928
step: 180, loss: 0.004649304319173098
step: 190, loss: 0.004436681978404522
step: 200, loss: 0.008347151800990105
step: 210, loss: 0.07175613939762115
step: 220, loss: 0.023261278867721558
step: 230, loss: 0.002608233131468296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9707207207207207, f1=0.9694915254237287,
best_f1=0.9694915254237287

```

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

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

```
step: 0, loss: 0.007210695184767246  
step: 10, loss: 0.05470436438918114  
step: 20, loss: 0.007489479146897793  
step: 30, loss: 0.009178660809993744  
step: 40, loss: 0.04261769726872444  
step: 50, loss: 0.004650669638067484  
step: 60, loss: 0.015803785994648933  
step: 70, loss: 0.05292782187461853  
step: 80, loss: 0.02373657375574112  
step: 90, loss: 0.04321132227778435  
step: 100, loss: 0.2484009563922882  
step: 110, loss: 0.004012362100183964  
step: 120, loss: 0.004729473497718573  
step: 130, loss: 0.005582293961197138  
step: 140, loss: 0.0038599807303398848  
step: 150, loss: 0.09781073778867722  
step: 160, loss: 0.008122995495796204  
step: 170, loss: 0.0029605396557599306  
step: 180, loss: 0.019199704751372337  
step: 190, loss: 0.020363081246614456  
step: 200, loss: 0.008443723432719707  
step: 210, loss: 0.0039291102439165115  
step: 220, loss: 0.05935647338628769  
step: 230, loss: 0.011352012865245342
```

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

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

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

```
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=0.9751693002257337, f1=0.9738933030646991,  
best_f1=0.9738933030646991
```

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

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

```
step: 0, loss: 0.008789975196123123  
step: 10, loss: 0.0034777484834194183  
step: 20, loss: 0.02192271500825882  
step: 30, loss: 0.002080244477838278  
step: 40, loss: 0.07408050447702408  
step: 50, loss: 0.04135587066411972  
step: 60, loss: 0.007531174924224615  
step: 70, loss: 0.061243291944265366  
step: 80, loss: 0.17156386375427246
```

```

step: 90, loss: 0.004615355748683214
step: 100, loss: 0.007225222419947386
step: 110, loss: 0.051832400262355804
step: 120, loss: 0.026252662762999535
step: 130, loss: 0.04039407894015312
step: 140, loss: 0.0028968113474547863
step: 150, loss: 0.01496000774204731
step: 160, loss: 0.0029385362286120653
step: 170, loss: 0.005288533866405487
step: 180, loss: 0.11404242366552353
step: 190, loss: 0.002205919474363327
step: 200, loss: 0.008438870310783386
step: 210, loss: 0.0027645668014883995
step: 220, loss: 0.0015094553818926215
step: 230, loss: 0.11845968663692474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9829738933030647, f1=0.9681093394077448,
best_f1=0.9681093394077448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009406639146618545
step: 10, loss: 0.003898950992152095
step: 20, loss: 0.0240839421749115
step: 30, loss: 0.00280473823659122
step: 40, loss: 0.0025137998163700104
step: 50, loss: 0.0017002613749355078
step: 60, loss: 0.016990534961223602
step: 70, loss: 0.04846176877617836
step: 80, loss: 0.025522420182824135
step: 90, loss: 0.08277846872806549
step: 100, loss: 0.0005810477887280285
step: 110, loss: 0.011251145042479038
step: 120, loss: 0.0009834780357778072
step: 130, loss: 0.0006948169320821762
step: 140, loss: 0.0020525252912193537
step: 150, loss: 0.012135711498558521
step: 160, loss: 0.0009284733678214252
step: 170, loss: 0.011683307588100433
step: 180, loss: 0.008187366649508476
step: 190, loss: 0.06327205151319504
step: 200, loss: 0.05988122150301933

```



```

step: 210, loss: 0.002086650114506483
step: 220, loss: 0.002598592545837164
step: 230, loss: 0.02415912039577961
/usr/local/lib/python3.7/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=0.9742441209406495, f1=0.9659090909090909,
best_f1=0.9681093394077448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011702129617333412
step: 10, loss: 0.0012484586331993341
step: 20, loss: 0.003940803464502096
step: 30, loss: 0.0008417824865318835
step: 40, loss: 0.00038511300226673484
step: 50, loss: 0.0015304672997444868
step: 60, loss: 0.0013709258055314422
step: 70, loss: 0.18502002954483032
step: 80, loss: 0.0015966787468641996
step: 90, loss: 0.01667633280158043
step: 100, loss: 0.004937690682709217
step: 110, loss: 0.04308806359767914
step: 120, loss: 0.0014726007357239723
step: 130, loss: 0.007435922976583242
step: 140, loss: 0.003126442665234208
step: 150, loss: 0.0007710673380643129
step: 160, loss: 0.0018064802279695868
step: 170, loss: 0.0029803693760186434
step: 180, loss: 0.0015793067868798971
step: 190, loss: 0.004401012789458036
step: 200, loss: 0.055431392043828964
step: 210, loss: 0.008189341984689236
step: 220, loss: 0.004100107587873936
step: 230, loss: 0.0010687824105843902
/usr/local/lib/python3.7/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=0.9811320754716982, f1=0.9721913236929923,
best_f1=0.9681093394077448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002364104613661766
step: 10, loss: 0.03227453678846359
step: 20, loss: 0.0045854756608605385
step: 30, loss: 0.00040186685509979725
step: 40, loss: 0.0026733623817563057

```

```

step: 50, loss: 0.0011137928813695908
step: 60, loss: 0.002574003068730235
step: 70, loss: 0.002027295995503664
step: 80, loss: 0.0006070288363844156
step: 90, loss: 0.008275202475488186
step: 100, loss: 0.0006421242724172771
step: 110, loss: 0.0021864157170057297
step: 120, loss: 0.0026429735589772463
step: 130, loss: 0.01874660700559616
step: 140, loss: 0.00036127344355918467
step: 150, loss: 0.1997772753238678
step: 160, loss: 0.0020357591565698385
step: 170, loss: 0.0922030434012413
step: 180, loss: 0.0005817920318804681
step: 190, loss: 0.0022831314709037542
step: 200, loss: 0.0861627385020256
step: 210, loss: 0.012272529304027557
step: 220, loss: 0.0006948741502128541
step: 230, loss: 0.0009227527189068496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830890642615557, f1=0.9736540664375716,
best_f1=0.9736540664375716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009936288697645068
step: 10, loss: 0.0017026507994160056
step: 20, loss: 0.004779871087521315
step: 30, loss: 0.001401441521011293
step: 40, loss: 0.0019396324642002583
step: 50, loss: 0.0018108203075826168
step: 60, loss: 0.0011127837933599949
step: 70, loss: 0.0005399337969720364
step: 80, loss: 0.0008253025007434189
step: 90, loss: 0.0007927344995550811
step: 100, loss: 0.0008639178122393787
step: 110, loss: 0.0031194586772471666
step: 120, loss: 0.0006649977876804769
step: 130, loss: 0.0066944388672709465
step: 140, loss: 0.0013007353991270065
step: 150, loss: 0.35433948040008545
step: 160, loss: 0.005643882788717747

```

```

step: 170, loss: 0.010841090232133865
step: 180, loss: 0.01223184447735548
step: 190, loss: 0.032053686678409576
step: 200, loss: 0.06814923882484436
step: 210, loss: 0.002026872243732214
step: 220, loss: 0.015825258567929268
step: 230, loss: 0.0005917919334024191
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9762174405436014,
best_f1=0.9736540664375716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006785758887417614
step: 10, loss: 0.001709267497062683
step: 20, loss: 0.0007214192301034927
step: 30, loss: 0.0010393073316663504
step: 40, loss: 0.006451253779232502
step: 50, loss: 0.000869586190674454
step: 60, loss: 0.000924373627640307
step: 70, loss: 0.02420814335346222
step: 80, loss: 0.0002250396355520934
step: 90, loss: 0.04422347992658615
step: 100, loss: 0.0011006038403138518
step: 110, loss: 0.00039309373823925853
step: 120, loss: 0.08197646588087082
step: 130, loss: 0.0009501442546024919
step: 140, loss: 0.007280828896909952
step: 150, loss: 0.001541727225022316
step: 160, loss: 0.001197846606373787
step: 170, loss: 0.0003705931594595313
step: 180, loss: 0.0024550631642341614
step: 190, loss: 0.0008872188627719879
step: 200, loss: 0.000624363252427429
step: 210, loss: 0.03300555422902107
step: 220, loss: 0.0015622712671756744
step: 230, loss: 0.0004909347044304013
/usr/local/lib/python3.7/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=0.9810055865921787, f1=0.9774774774774775,
best_f1=0.9736540664375716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033461517887189984

```

```

step: 10, loss: 0.0004489355778787285
step: 20, loss: 0.0008752605644986033
step: 30, loss: 0.0036219183821231127
step: 40, loss: 0.001672547310590744
step: 50, loss: 0.001153496908955276
step: 60, loss: 0.0006544497446157038
step: 70, loss: 0.007746197283267975
step: 80, loss: 0.00023612617223989218
step: 90, loss: 0.00029979966348037124
step: 100, loss: 0.0005021481192670763
step: 110, loss: 0.00035448870039545
step: 120, loss: 0.00021772868058178574
step: 130, loss: 0.0010971324518322945
step: 140, loss: 0.0002565820177551359
step: 150, loss: 0.0022077267058193684
step: 160, loss: 0.00014577136607840657
step: 170, loss: 0.00041598716052249074
step: 180, loss: 0.0004119670484215021
step: 190, loss: 0.0008118175901472569
step: 200, loss: 0.0027367600705474615
step: 210, loss: 0.0004645981825888157
step: 220, loss: 0.015775755047798157
step: 230, loss: 0.00041420268826186657
/usr/local/lib/python3.7/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=0.978912319644839, f1=0.9755555555555556,
best_f1=0.9736540664375716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000231456637266092
step: 10, loss: 0.023512395098805428
step: 20, loss: 0.0005694996798411012
step: 30, loss: 0.00025385990738868713
step: 40, loss: 0.00012465196778066456
step: 50, loss: 0.00016720563871785998
step: 60, loss: 0.009083153679966927
step: 70, loss: 0.002738287905231118
step: 80, loss: 0.0013566690031439066
step: 90, loss: 0.1785934716463089
step: 100, loss: 0.0007412764825858176
step: 110, loss: 0.0051676868461072445
step: 120, loss: 0.0004066030087415129
step: 130, loss: 0.0006762047414667904
step: 140, loss: 0.0027024727314710617
step: 150, loss: 0.0006017806008458138
step: 160, loss: 0.0005183489993214607

```

```

step: 170, loss: 0.008245397359132767
step: 180, loss: 0.0003291405737400055
step: 190, loss: 0.0003434818936511874
step: 200, loss: 0.0011383468518033624
step: 210, loss: 0.0006285022245720029
step: 220, loss: 0.0007268537301570177
step: 230, loss: 0.00048370883450843394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9820224719101124, f1=0.9739524348810873,
best_f1=0.9736540664375716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023747205268591642
step: 10, loss: 0.0004751123196911067
step: 20, loss: 0.03823917359113693
step: 30, loss: 0.044657301157712936
step: 40, loss: 0.0008954777149483562
step: 50, loss: 0.002079547382891178
step: 60, loss: 0.0008847851422615349
step: 70, loss: 0.00023465321282856166
step: 80, loss: 0.0001387542433803901
step: 90, loss: 0.00048367466661147773
step: 100, loss: 0.0001337618741672486
step: 110, loss: 0.00015064810577314347
step: 120, loss: 0.0002654869749676436
step: 130, loss: 0.00019048582180403173
step: 140, loss: 0.0007946338737383485
step: 150, loss: 0.0005867391591891646
step: 160, loss: 0.0005653816042467952
step: 170, loss: 0.00041515176417306066
step: 180, loss: 0.002587389200925827
step: 190, loss: 0.006012956146150827
step: 200, loss: 0.00020815987954847515
step: 210, loss: 0.09958595037460327
step: 220, loss: 0.019022837281227112
step: 230, loss: 0.009958979673683643
/usr/local/lib/python3.7/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=0.98, f1=0.9712389380530975, best_f1=0.9736540664375716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00047632999485358596
step: 10, loss: 0.005839257501065731

```

```

step: 20, loss: 0.0019463833887130022
step: 30, loss: 0.00014835354522801936
step: 40, loss: 0.0014289289247244596
step: 50, loss: 0.0009127368684858084
step: 60, loss: 0.00022455822909250855
step: 70, loss: 0.000588360067922622
step: 80, loss: 0.000370017223758623
step: 90, loss: 0.0003722166584338993
step: 100, loss: 0.001521353144198656
step: 110, loss: 0.0011290202382951975
step: 120, loss: 0.0033208809327334166
step: 130, loss: 0.00034540731576271355
step: 140, loss: 0.00032065852428786457
step: 150, loss: 0.00018104717310052365
step: 160, loss: 0.001057485118508339
step: 170, loss: 0.0005812563467770815
step: 180, loss: 0.03544577583670616
step: 190, loss: 0.02342817559838295
step: 200, loss: 9.944449993781745e-05
step: 210, loss: 0.0007062137010507286
step: 220, loss: 0.0004810462996829301
step: 230, loss: 0.0012766100699082017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, 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: 0.00019305052410345525
step: 10, loss: 0.00019875327416229993
step: 20, loss: 0.0005177344428375363
step: 30, loss: 0.004655186086893082
step: 40, loss: 0.0003223624371457845
step: 50, loss: 0.0002678789896890521
step: 60, loss: 0.00037459711893461645
step: 70, loss: 0.00022071179409977049
step: 80, loss: 0.0003112855483777821
step: 90, loss: 0.0027960999868810177
step: 100, loss: 0.0003938625450246036
step: 110, loss: 0.00046124885557219386
step: 120, loss: 7.232677307911217e-05
step: 130, loss: 0.0032938257791101933

```

```

step: 140, loss: 0.00032818742329254746
step: 150, loss: 0.00022511344286613166
step: 160, loss: 0.00034435122506693006
step: 170, loss: 0.0003434046229813248
step: 180, loss: 0.012684782035648823
step: 190, loss: 0.0001979100052267313
step: 200, loss: 0.00031055163708515465
step: 210, loss: 0.00017750148253981024
step: 220, loss: 0.0002963655279017985
step: 230, loss: 0.003089002100750804
/usr/local/lib/python3.7/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=0.9810901001112348, f1=0.9765363128491621,
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: 0.005262135993689299
step: 10, loss: 0.00018283419194631279
step: 20, loss: 0.00020964523719158024
step: 30, loss: 0.0002458170347381383
step: 40, loss: 0.00012797310773748904
step: 50, loss: 0.00023314192367251962
step: 60, loss: 0.021303851157426834
step: 70, loss: 0.0006139411125332117
step: 80, loss: 0.00018619251204654574
step: 90, loss: 0.00014607131015509367
step: 100, loss: 0.00010666863818187267
step: 110, loss: 0.014011364430189133
step: 120, loss: 0.0484914593398571
step: 130, loss: 0.0002749269478954375
step: 140, loss: 0.012401510030031204
step: 150, loss: 0.02153814397752285
step: 160, loss: 0.007220679894089699
step: 170, loss: 7.697525870753452e-05
step: 180, loss: 0.0006110559916123748
step: 190, loss: 0.0020981300622224808
step: 200, loss: 0.001749483635649085
step: 210, loss: 0.03325942903757095
step: 220, loss: 0.0001833968417486176
step: 230, loss: 0.0006082928739488125
/usr/local/lib/python3.7/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=0.9831649831649831, f1=0.9808773903262092,
best_f1=0.9763779527559054
mv: cannot stat '*_dev.pt': No such file or directory

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

2473it [00:12, 204.74it/s]

load_f1 = 0.9876265466816648

real_f1 = 0.9876265466816648

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

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

4398it [00:22, 195.03it/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/

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6389181613922119
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: 0.4807523488998413
step: 20, loss: 0.29640886187553406
step: 30, loss: 0.36514610052108765
step: 40, loss: 0.37766778469085693
step: 50, loss: 0.44229304790496826
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.29931437969207764
step: 70, loss: 0.3358169198036194
step: 80, loss: 0.18196380138397217
step: 90, loss: 0.14045250415802002
step: 100, loss: 0.2517932057380676
step: 110, loss: 0.10763632506132126
step: 120, loss: 0.16902022063732147
step: 130, loss: 0.09746725857257843
step: 140, loss: 0.2032417356967926
step: 150, loss: 0.0866924524307251
step: 160, loss: 0.1406531035900116
step: 170, loss: 0.3614106774330139
step: 180, loss: 0.03133748844265938
step: 190, loss: 0.09072252362966537
step: 200, loss: 0.04405752941966057
step: 210, loss: 0.08754287660121918
step: 220, loss: 0.05940881371498108
step: 230, loss: 0.11850877106189728
step: 240, loss: 0.031420737504959106
step: 250, loss: 0.05061115697026253
step: 260, loss: 0.06046416983008385
step: 270, loss: 0.21674136817455292
step: 280, loss: 0.04900047183036804
step: 290, loss: 0.05547222122550011
step: 300, loss: 0.19200773537158966
step: 310, loss: 0.08708652853965759
step: 320, loss: 0.10890824347734451
step: 330, loss: 0.05610084906220436
step: 340, loss: 0.3128112256526947
step: 350, loss: 0.08584794402122498
step: 360, loss: 0.06258148699998856
step: 370, loss: 0.01701820269227028
step: 380, loss: 0.11887002736330032

```

```

step: 390, loss: 0.04652411490678787
step: 400, loss: 0.04998067021369934
step: 410, loss: 0.28006675839424133
step: 420, loss: 0.03517843037843704
step: 430, loss: 0.04912026599049568
step: 440, loss: 0.06319133192300797
step: 450, loss: 0.10163194686174393
step: 460, loss: 0.005902111530303955
step: 470, loss: 0.06731579452753067
step: 480, loss: 0.1670626997947693
step: 490, loss: 0.04337061941623688
step: 500, loss: 0.013286575675010681
step: 510, loss: 0.02057904563844204
step: 520, loss: 0.03842976316809654
step: 530, loss: 0.019828105345368385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9487648673376029, f1=0.9445716903344022,
best_f1=0.9445716903344022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13271625339984894
step: 10, loss: 0.06355747580528259
step: 20, loss: 0.06666363775730133
step: 30, loss: 0.04223538562655449
step: 40, loss: 0.07706977427005768
step: 50, loss: 0.06137111783027649
step: 60, loss: 0.04283139854669571
step: 70, loss: 0.02289758436381817
step: 80, loss: 0.01382574625313282
step: 90, loss: 0.018370578065514565
step: 100, loss: 0.16020677983760834
step: 110, loss: 0.007732793688774109
step: 120, loss: 0.01529358234256506
step: 130, loss: 0.01289287954568863
step: 140, loss: 0.05847031623125076
step: 150, loss: 0.04296698048710823
step: 160, loss: 0.049336690455675125
step: 170, loss: 0.03697465732693672
step: 180, loss: 0.07323967665433884
step: 190, loss: 0.021332601085305214
step: 200, loss: 0.2019825428724289

```

```

step: 210, loss: 0.019388940185308456
step: 220, loss: 0.0008661414030939341
step: 230, loss: 0.07334735989570618
step: 240, loss: 0.11999582499265671
step: 250, loss: 0.03643534705042839
step: 260, loss: 0.04452981799840927
step: 270, loss: 0.13629940152168274
step: 280, loss: 0.05401180684566498
step: 290, loss: 0.07240594923496246
step: 300, loss: 0.07656314224004745
step: 310, loss: 0.07381592690944672
step: 320, loss: 0.07153624296188354
step: 330, loss: 0.05095726624131203
step: 340, loss: 0.1019919142127037
step: 350, loss: 0.0029634973034262657
step: 360, loss: 0.10403122007846832
step: 370, loss: 0.014505824074149132
step: 380, loss: 0.1448758840560913
step: 390, loss: 0.002967034000903368
step: 400, loss: 0.045721858739852905
step: 410, loss: 0.07865604758262634
step: 420, loss: 0.043166473507881165
step: 430, loss: 0.3405137062072754
step: 440, loss: 0.013604076579213142
step: 450, loss: 0.04891883581876755
step: 460, loss: 0.059305042028427124
step: 470, loss: 0.051645196974277496
step: 480, loss: 0.045213669538497925
step: 490, loss: 0.0518772155046463
step: 500, loss: 0.0074969688430428505
step: 510, loss: 0.005888830404728651
step: 520, loss: 0.3626055121421814
step: 530, loss: 0.10586081445217133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9522935779816514, f1=0.9465020576131687,
best_f1=0.9465020576131687
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15579776465892792
step: 10, loss: 0.021681521087884903
step: 20, loss: 0.014593447558581829

```

step: 30, loss: 0.02507769502699375
step: 40, loss: 0.038937170058488846
step: 50, loss: 0.02357572130858898
step: 60, loss: 0.10037875920534134
step: 70, loss: 0.014742448925971985
step: 80, loss: 0.03696851432323456
step: 90, loss: 0.04147431626915932
step: 100, loss: 0.020608384162187576
step: 110, loss: 0.01053224690258503
step: 120, loss: 0.330801397562027
step: 130, loss: 0.07744099944829941
step: 140, loss: 0.015829112380743027
step: 150, loss: 0.02337600290775299
step: 160, loss: 0.04365048557519913
step: 170, loss: 0.0018323957920074463
step: 180, loss: 0.04819362610578537
step: 190, loss: 0.0016090284334495664
step: 200, loss: 0.05325644090771675
step: 210, loss: 0.031177490949630737
step: 220, loss: 0.08433473855257034
step: 230, loss: 0.12283247709274292
step: 240, loss: 0.026245059445500374
step: 250, loss: 0.046861838549375534
step: 260, loss: 0.16729891300201416
step: 270, loss: 0.011245505884289742
step: 280, loss: 0.06547068804502487
step: 290, loss: 0.021554559469223022
step: 300, loss: 0.0690436065196991
step: 310, loss: 0.06759177148342133
step: 320, loss: 0.031522952020168304
step: 330, loss: 0.0124617088586092
step: 340, loss: 0.014697586186230183
step: 350, loss: 0.18512137234210968
step: 360, loss: 0.026515142992138863
step: 370, loss: 0.03556329384446144
step: 380, loss: 0.0026335474103689194
step: 390, loss: 0.006546803284436464
step: 400, loss: 0.10993879288434982
step: 410, loss: 0.051731932908296585
step: 420, loss: 0.023199981078505516
step: 430, loss: 0.024900807067751884
step: 440, loss: 0.13109815120697021
step: 450, loss: 0.0507875457406044
step: 460, loss: 0.10083805024623871
step: 470, loss: 0.007017447147518396
step: 480, loss: 0.11147351562976837
step: 490, loss: 0.04300592839717865
step: 500, loss: 0.015273363329470158

```

step: 510, loss: 0.028732793405652046
step: 520, loss: 0.007984189316630363
step: 530, loss: 0.0026859953068196774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9592592592592593, f1=0.9506057781919852,
best_f1=0.9506057781919852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0244881734251976
step: 10, loss: 0.0014163688756525517
step: 20, loss: 0.0080211590975523
step: 30, loss: 0.15611742436885834
step: 40, loss: 0.002638560486957431
step: 50, loss: 0.024066150188446045
step: 60, loss: 0.015500354580581188
step: 70, loss: 0.05396364629268646
step: 80, loss: 0.035807427018880844
step: 90, loss: 0.1351979374885559
step: 100, loss: 0.016019051894545555
step: 110, loss: 0.1231193020939827
step: 120, loss: 0.0019151101587340236
step: 130, loss: 0.1366480588912964
step: 140, loss: 0.027820613235235214
step: 150, loss: 0.007839696481823921
step: 160, loss: 0.0048653241246938705
step: 170, loss: 0.009627123363316059
step: 180, loss: 0.09671429544687271
step: 190, loss: 0.01725323684513569
step: 200, loss: 0.010090198367834091
step: 210, loss: 0.00783065240830183
step: 220, loss: 0.0291278213262558
step: 230, loss: 0.019893193617463112
step: 240, loss: 0.003592569613829255
step: 250, loss: 0.06510469317436218
step: 260, loss: 0.02403385192155838
step: 270, loss: 0.09799898415803909
step: 280, loss: 0.00554202776402235
step: 290, loss: 0.024648847058415413
step: 300, loss: 0.005935213528573513
step: 310, loss: 0.003415332641452551
step: 320, loss: 0.083197221159935

```

```

step: 330, loss: 0.10550115257501602
step: 340, loss: 0.005334951914846897
step: 350, loss: 0.008632350713014603
step: 360, loss: 0.06605248153209686
step: 370, loss: 0.002978381933644414
step: 380, loss: 0.021839933469891548
step: 390, loss: 0.0003285334096290171
step: 400, loss: 0.028171777725219727
step: 410, loss: 0.09205958992242813
step: 420, loss: 0.015968037769198418
step: 430, loss: 0.01370697095990181
step: 440, loss: 0.04921974614262581
step: 450, loss: 0.0198118444532156
step: 460, loss: 0.010354118421673775
step: 470, loss: 0.0015339504461735487
step: 480, loss: 0.0016758308047428727
step: 490, loss: 0.0011439428199082613
step: 500, loss: 0.054381098598241806
step: 510, loss: 0.13294337689876556
step: 520, loss: 0.03161279857158661
step: 530, loss: 0.13769201934337616
/usr/local/lib/python3.7/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=0.9579676674364896, f1=0.950485886163813,
best_f1=0.9506057781919852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0055802264250814915
step: 10, loss: 0.006883231922984123
step: 20, loss: 0.029215008020401
step: 30, loss: 0.03138405457139015
step: 40, loss: 0.009928898885846138
step: 50, loss: 0.04086754098534584
step: 60, loss: 0.03039882332086563
step: 70, loss: 0.0012444909662008286
step: 80, loss: 0.0020109128672629595
step: 90, loss: 0.03829793259501457
step: 100, loss: 0.02995847538113594
step: 110, loss: 0.005487052723765373
step: 120, loss: 0.10290934145450592
step: 130, loss: 0.0027725494001060724
step: 140, loss: 0.03457709774374962
step: 150, loss: 0.005083841271698475
step: 160, loss: 0.013562728650867939
step: 170, loss: 0.0538848377764225
step: 180, loss: 0.01491823885589838

```

```

step: 190, loss: 0.019528765231370926
step: 200, loss: 0.009509778581559658
step: 210, loss: 0.0006403069710358977
step: 220, loss: 0.0031388297211378813
step: 230, loss: 0.0009438595734536648
step: 240, loss: 0.032095734030008316
step: 250, loss: 0.12355603277683258
step: 260, loss: 0.0029449493158608675
step: 270, loss: 0.013438272289931774
step: 280, loss: 0.00450926274061203
step: 290, loss: 0.002958059310913086
step: 300, loss: 0.05119618773460388
step: 310, loss: 0.06436828523874283
step: 320, loss: 0.11134273558855057
step: 330, loss: 0.0002765884855762124
step: 340, loss: 0.02582216076552868
step: 350, loss: 0.0027400259859859943
step: 360, loss: 0.003986597526818514
step: 370, loss: 0.0007198986131697893
step: 380, loss: 0.006896481849253178
step: 390, loss: 0.009217249229550362
step: 400, loss: 0.02701755240559578
step: 410, loss: 0.015258881263434887
step: 420, loss: 0.21111652255058289
step: 430, loss: 0.0031614964827895164
step: 440, loss: 0.008302474394440651
step: 450, loss: 0.0056404126808047295
step: 460, loss: 0.02182554453611374
step: 470, loss: 0.02643604204058647
step: 480, loss: 0.06808944046497345
step: 490, loss: 0.008985921740531921
step: 500, loss: 0.03999384492635727
step: 510, loss: 0.019093016162514687
step: 520, loss: 0.06111007556319237
step: 530, loss: 0.03889232501387596
/usr/local/lib/python3.7/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=0.955113373438223, f1=0.945845004668534,
best_f1=0.9506057781919852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011216132901608944
step: 10, loss: 0.0327012799680233
step: 20, loss: 0.0013837066944688559
step: 30, loss: 0.00110229616984725
step: 40, loss: 0.0005253879353404045

```


step: 50, loss: 0.0077755749225616455
step: 60, loss: 0.003087090328335762
step: 70, loss: 0.0013453762512654066
step: 80, loss: 0.0007769197691231966
step: 90, loss: 0.001193263684399426
step: 100, loss: 0.01112572755664587
step: 110, loss: 0.01581924967467785
step: 120, loss: 0.03352367877960205
step: 130, loss: 0.008017381653189659
step: 140, loss: 0.005537023767828941
step: 150, loss: 0.010620282031595707
step: 160, loss: 0.0488785021007061
step: 170, loss: 0.0014503082493320107
step: 180, loss: 0.006761699449270964
step: 190, loss: 0.03456120938062668
step: 200, loss: 0.003980178385972977
step: 210, loss: 0.006604727823287249
step: 220, loss: 0.007059136405587196
step: 230, loss: 0.0338028222322464
step: 240, loss: 0.02078883908689022
step: 250, loss: 0.015274852514266968
step: 260, loss: 0.008122487924993038
step: 270, loss: 0.00059311039512977
step: 280, loss: 0.0007941558142192662
step: 290, loss: 0.00020793767180293798
step: 300, loss: 0.028400011360645294
step: 310, loss: 0.13560806214809418
step: 320, loss: 0.0014416532358154655
step: 330, loss: 0.010177606716752052
step: 340, loss: 0.0007421618211083114
step: 350, loss: 0.0007424066425301135
step: 360, loss: 0.0320446752011776
step: 370, loss: 0.010068653151392937
step: 380, loss: 0.0011334443697705865
step: 390, loss: 0.005265609361231327
step: 400, loss: 0.011937825009226799
step: 410, loss: 0.0011164001189172268
step: 420, loss: 0.0036070269998162985
step: 430, loss: 0.0005849471199326217
step: 440, loss: 0.0009758686646819115
step: 450, loss: 0.2186802327632904
step: 460, loss: 0.007222642190754414
step: 470, loss: 0.0037200981751084328
step: 480, loss: 0.0037186474073678255
step: 490, loss: 0.0028238107915967703
step: 500, loss: 0.001885687350295484
step: 510, loss: 0.10294012725353241
step: 520, loss: 0.0010394547134637833

```
step: 530, loss: 0.06039899215102196
/usr/local/lib/python3.7/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=0.9553903345724908, f1=0.9454036397573495,
best_f1=0.9506057781919852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004421710502356291
step: 10, loss: 0.003397090593352914
step: 20, loss: 0.0007779565639793873
step: 30, loss: 0.007383196614682674
step: 40, loss: 0.039156194776296616
step: 50, loss: 0.008982122875750065
step: 60, loss: 0.0039007430896162987
step: 70, loss: 0.025947656482458115
step: 80, loss: 0.00045534735545516014
step: 90, loss: 0.00017114182992372662
step: 100, loss: 0.023167584091424942
step: 110, loss: 0.0006485944031737745
step: 120, loss: 0.08094184100627899
step: 130, loss: 0.000566133123356849
step: 140, loss: 0.003643654054030776
step: 150, loss: 0.0023822051007300615
step: 160, loss: 0.015634827315807343
step: 170, loss: 0.02279786206781864
step: 180, loss: 0.025080937892198563
step: 190, loss: 0.01695682480931282
step: 200, loss: 3.8268590287771076e-05
step: 210, loss: 0.002951947972178459
step: 220, loss: 0.00045742763904854655
step: 230, loss: 0.0018538528820499778
step: 240, loss: 0.04971475154161453
step: 250, loss: 0.016917988657951355
step: 260, loss: 0.006279681343585253
step: 270, loss: 0.004550079349428415
step: 280, loss: 0.007451205980032682
step: 290, loss: 0.023776069283485413
step: 300, loss: 0.0001556901406729594
step: 310, loss: 0.0008885699789971113
step: 320, loss: 0.005495184101164341
step: 330, loss: 6.76168201607652e-05
step: 340, loss: 0.01114397868514061
step: 350, loss: 0.00389049737714231
step: 360, loss: 0.014564494602382183
step: 370, loss: 0.006108424626290798
step: 380, loss: 0.0360863022506237
```

```

step: 390, loss: 0.004860400687903166
step: 400, loss: 0.03666064515709877
step: 410, loss: 0.012153558433055878
step: 420, loss: 0.003309921594336629
step: 430, loss: 0.01846572384238243
step: 440, loss: 0.005839843302965164
step: 450, loss: 0.0003923306649085134
step: 460, loss: 0.0014045336283743382
step: 470, loss: 0.1518971174955368
step: 480, loss: 0.002296310616657138
step: 490, loss: 0.0032927622087299824
step: 500, loss: 0.00294682988896966
step: 510, loss: 0.0008702545310370624
step: 520, loss: 0.0001932816085172817
step: 530, loss: 0.009658509865403175
/usr/local/lib/python3.7/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=0.9575871819038644, f1=0.9427895981087471,
best_f1=0.9506057781919852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00049120670882985
step: 10, loss: 0.001497686724178493
step: 20, loss: 0.057921335101127625
step: 30, loss: 0.03772743418812752
step: 40, loss: 0.0001684423186816275
step: 50, loss: 0.0015962880570441484
step: 60, loss: 0.0013702823780477047
step: 70, loss: 0.02741824835538864
step: 80, loss: 0.030244028195738792
step: 90, loss: 0.00545905064791441
step: 100, loss: 0.0008983754669316113
step: 110, loss: 0.002563341287896037
step: 120, loss: 0.0012532612308859825
step: 130, loss: 0.006202720571309328
step: 140, loss: 0.003987433388829231
step: 150, loss: 0.007643776014447212
step: 160, loss: 0.00245493883267045
step: 170, loss: 0.11934980005025864
step: 180, loss: 0.005727660376578569
step: 190, loss: 0.013550440780818462
step: 200, loss: 0.002699099248275161
step: 210, loss: 0.06458007544279099
step: 220, loss: 0.0008317831088788807
step: 230, loss: 0.05955575406551361
step: 240, loss: 0.1411542296409607

```

```

step: 250, loss: 0.00991777703166008
step: 260, loss: 0.014207689091563225
step: 270, loss: 0.009907295927405357
step: 280, loss: 0.000792577862739563
step: 290, loss: 0.004349098540842533
step: 300, loss: 9.980559116229415e-05
step: 310, loss: 0.0013882501516491175
step: 320, loss: 0.04384470358490944
step: 330, loss: 0.00016048412362579256
step: 340, loss: 0.014865336939692497
step: 350, loss: 9.226218389812857e-05
step: 360, loss: 0.013078445568680763
step: 370, loss: 0.027591120451688766
step: 380, loss: 0.0005289735272526741
step: 390, loss: 0.007923219352960587
step: 400, loss: 0.00773041183128953
step: 410, loss: 0.00046608634875155985
step: 420, loss: 0.0003296244831290096
step: 430, loss: 0.018850574269890785
step: 440, loss: 0.013349037617444992
step: 450, loss: 0.00013841538748238236
step: 460, loss: 0.004147314000874758
step: 470, loss: 0.02202555350959301
step: 480, loss: 0.002257764572277665
step: 490, loss: 0.000847310817334801
step: 500, loss: 0.00024809889146126807
step: 510, loss: 0.001871449057944119
step: 520, loss: 0.00019599890219978988
step: 530, loss: 0.0013677289243787527
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9568014705882353, f1=0.9479981592268752,
best_f1=0.9506057781919852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014013935287948698
step: 10, loss: 0.015094572678208351
step: 20, loss: 0.0004738125717267394
step: 30, loss: 0.08484449237585068
step: 40, loss: 0.0001277923001907766
step: 50, loss: 5.227440851740539e-05
step: 60, loss: 0.00021078623831272125
step: 70, loss: 0.025230038911104202
step: 80, loss: 0.01974235288798809
step: 90, loss: 0.0660879909992218

```

```
step: 100, loss: 0.00018093411927111447
step: 110, loss: 0.0011305399239063263
step: 120, loss: 0.005723552778363228
step: 130, loss: 0.006048902869224548
step: 140, loss: 0.004323569126427174
step: 150, loss: 0.0004910108400508761
step: 160, loss: 0.0007477265899069607
step: 170, loss: 0.01698015071451664
step: 180, loss: 0.00045938545372337103
step: 190, loss: 0.007087864447385073
step: 200, loss: 0.00024865398881956935
step: 210, loss: 0.0003605485544539988
step: 220, loss: 0.00046256702626124024
step: 230, loss: 0.0002980790741275996
step: 240, loss: 0.00023489865998271853
step: 250, loss: 0.00035932910395786166
step: 260, loss: 0.0004936076002195477
step: 270, loss: 0.025204092264175415
step: 280, loss: 0.014780937694013119
step: 290, loss: 0.0010804791236296296
step: 300, loss: 0.0007267709588631988
step: 310, loss: 0.0006839397829025984
step: 320, loss: 0.0005861423560418189
step: 330, loss: 0.0006783127319067717
step: 340, loss: 0.0053163859993219376
step: 350, loss: 0.07998625934123993
step: 360, loss: 0.0037853082176297903
step: 370, loss: 0.00014930963516235352
step: 380, loss: 0.0003539708850439638
step: 390, loss: 5.76038692088332e-05
step: 400, loss: 0.01122453436255455
step: 410, loss: 0.001318734372034669
step: 420, loss: 0.0004859082691837102
step: 430, loss: 0.004071306902915239
step: 440, loss: 9.33730261749588e-05
step: 450, loss: 0.00026040771626867354
step: 460, loss: 0.00034206616692245007
step: 470, loss: 0.00036648361128754914
step: 480, loss: 0.0005423488328233361
step: 490, loss: 0.03683335334062576
step: 500, loss: 0.000964783423114568
step: 510, loss: 0.0007288750493898988
step: 520, loss: 0.006332279182970524
step: 530, loss: 0.038994379341602325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9618604651162791, f1=0.946927374301676,

best_f1=0.946927374301676

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

An input tensor was not cuda.

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

step: 0, loss: 0.00173133984208107

step: 10, loss: 0.001876911148428917

step: 20, loss: 0.00018910686776507646

step: 30, loss: 0.0034360529389232397

step: 40, loss: 0.003876002738252282

step: 50, loss: 0.0011126534081995487

step: 60, loss: 0.010935296304523945

step: 70, loss: 0.00020705681527033448

step: 80, loss: 0.018075890839099884

step: 90, loss: 0.0004539231595117599

step: 100, loss: 0.000597361009567976

step: 110, loss: 0.0025532275903970003

step: 120, loss: 0.0005335621535778046

step: 130, loss: 0.0005057416274212301

step: 140, loss: 0.0008268014644272625

step: 150, loss: 0.00028402142925187945

step: 160, loss: 0.03930041566491127

step: 170, loss: 0.00017729429237078875

step: 180, loss: 0.012868385761976242

step: 190, loss: 0.0013362972531467676

step: 200, loss: 0.00846664048731327

step: 210, loss: 0.012507640756666666

step: 220, loss: 0.0011979836272075772

step: 230, loss: 0.0003875381371472031

step: 240, loss: 0.0028815034311264753

step: 250, loss: 0.0013998269569128752

step: 260, loss: 0.000286743335891515

step: 270, loss: 5.992344085825607e-05

step: 280, loss: 0.03325723484158516

step: 290, loss: 0.00026067456929013133

step: 300, loss: 0.008331025950610638

step: 310, loss: 0.16096536815166473

step: 320, loss: 0.07696511596441269

step: 330, loss: 0.008956694975495338

step: 340, loss: 0.007759135216474533

step: 350, loss: 0.0018698071362450719

step: 360, loss: 0.001996389590203762

step: 370, loss: 0.0005403751856647432

step: 380, loss: 0.00244667986407876

step: 390, loss: 8.120207348838449e-05

```

step: 400, loss: 0.002594125922769308
step: 410, loss: 0.0025873067788779736
step: 420, loss: 0.0009583774954080582
step: 430, loss: 0.0024321957025676966
step: 440, loss: 0.0003409223281778395
step: 450, loss: 0.006422318052500486
step: 460, loss: 0.0026795011945068836
step: 470, loss: 0.0013350893277674913
step: 480, loss: 0.00036883013672195375
step: 490, loss: 0.00023908872390165925
step: 500, loss: 0.0008429201552644372
step: 510, loss: 0.00012118802987970412
step: 520, loss: 0.024946438148617744
step: 530, loss: 0.009556131437420845
/usr/local/lib/python3.7/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=0.9549295774647887, f1=0.9455909943714823,
best_f1=0.946927374301676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002881482942029834
step: 10, loss: 0.21110652387142181
step: 20, loss: 0.0010727185290306807
step: 30, loss: 0.0028409354854375124
step: 40, loss: 0.00016490639245603234
step: 50, loss: 7.245615415740758e-05
step: 60, loss: 0.0003569349937606603
step: 70, loss: 0.0001682090514805168
step: 80, loss: 0.00364185543730855
step: 90, loss: 0.005917079281061888
step: 100, loss: 0.013443011790513992
step: 110, loss: 0.0007016423624008894
step: 120, loss: 0.008294633589684963
step: 130, loss: 0.00033741333754733205
step: 140, loss: 0.0008589706849306822
step: 150, loss: 0.0022406261414289474
step: 160, loss: 0.0009001697180792689
step: 170, loss: 0.0015587519155815244
step: 180, loss: 0.0013686129823327065
step: 190, loss: 0.048043232411146164
step: 200, loss: 0.03565787151455879
step: 210, loss: 0.0013597388751804829
step: 220, loss: 0.00676722964271903
step: 230, loss: 0.0007267020992003381
step: 240, loss: 0.018443087115883827
step: 250, loss: 8.877259097062051e-05

```

```

step: 260, loss: 0.0012474972754716873
step: 270, loss: 0.003886465448886156
step: 280, loss: 0.0011406611884012818
step: 290, loss: 0.005757385864853859
step: 300, loss: 0.0015266984701156616
step: 310, loss: 0.03351544961333275
step: 320, loss: 0.0039765918627381325
step: 330, loss: 0.00037730939220637083
step: 340, loss: 0.007105004508048296
step: 350, loss: 0.0033563554752618074
step: 360, loss: 0.0005916391965001822
step: 370, loss: 0.002699383068829775
step: 380, loss: 0.000277179351542145
step: 390, loss: 0.0013346350751817226
step: 400, loss: 0.00010892922000493854
step: 410, loss: 0.001420897082425654
step: 420, loss: 0.0041457246989011765
step: 430, loss: 0.0023460248485207558
step: 440, loss: 0.0002791738079395145
step: 450, loss: 0.0008312942227348685
step: 460, loss: 0.004150954075157642
step: 470, loss: 9.667244012234733e-05
step: 480, loss: 0.00027026014868170023
step: 490, loss: 0.14323194324970245
step: 500, loss: 0.010131286457180977
step: 510, loss: 0.020542927086353302
step: 520, loss: 0.01571817137300968
step: 530, loss: 0.011464369483292103
/usr/local/lib/python3.7/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=0.9562010142923005, f1=0.943239501615136,
best_f1=0.946927374301676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015648540284018964
step: 10, loss: 0.01023293286561966
step: 20, loss: 0.003315693698823452
step: 30, loss: 3.143255162285641e-05
step: 40, loss: 0.00042940492858178914
step: 50, loss: 0.009153547696769238
step: 60, loss: 0.0009009091299958527
step: 70, loss: 0.028090085834264755
step: 80, loss: 0.00018618078320287168
step: 90, loss: 0.026798248291015625
step: 100, loss: 0.010978253558278084
step: 110, loss: 0.01968657411634922

```



```

step: 120, loss: 0.01088695228099823
step: 130, loss: 0.0006210222491063178
step: 140, loss: 0.00015444702876266092
step: 150, loss: 0.0004204268625471741
step: 160, loss: 0.0005377838388085365
step: 170, loss: 0.00020766454690601677
step: 180, loss: 0.00011528535833349451
step: 190, loss: 0.00029025549883954227
step: 200, loss: 0.0014850451843813062
step: 210, loss: 0.0006341454572975636
step: 220, loss: 0.0010205009020864964
step: 230, loss: 0.00015425161109305918
step: 240, loss: 0.0009333559428341687
step: 250, loss: 2.956613025162369e-05
step: 260, loss: 0.0005828404100611806
step: 270, loss: 0.0006403890438377857
step: 280, loss: 5.431898171082139e-05
step: 290, loss: 0.00016822559700813144
step: 300, loss: 0.006641800981014967
step: 310, loss: 0.0014611922670155764
step: 320, loss: 0.00044575153151527047
step: 330, loss: 0.01988321542739868
step: 340, loss: 0.00017919845413416624
step: 350, loss: 5.431584577308968e-05
step: 360, loss: 0.007478204555809498
step: 370, loss: 0.012469960376620293
step: 380, loss: 0.0003856909170281142
step: 390, loss: 0.002709614345803857
step: 400, loss: 6.863632006570697e-05
step: 410, loss: 0.00021276077313814312
step: 420, loss: 0.016441501677036285
step: 430, loss: 0.00029906531563028693
step: 440, loss: 0.00401055533438921
step: 450, loss: 0.010967257432639599
step: 460, loss: 0.0004972020396962762
step: 470, loss: 0.006846357136964798
step: 480, loss: 0.08962947875261307
step: 490, loss: 0.008866848424077034
step: 500, loss: 6.859389395685866e-05
step: 510, loss: 0.043952684849500656
step: 520, loss: 0.0025993504095822573
step: 530, loss: 0.012450107373297215
/usr/local/lib/python3.7/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=0.9557933922754769, f1=0.9499536607970344,
best_f1=0.946927374301676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

step: 0, loss: 0.0003139445907436311

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

step: 10, loss: 0.00021199713228270411

step: 20, loss: 0.003453125711530447

step: 30, loss: 2.4891371140256524e-05

step: 40, loss: 0.00020163372391834855

step: 50, loss: 0.0010728383203968406

step: 60, loss: 0.0005034275236539543

step: 70, loss: 0.0024304380640387535

step: 80, loss: 0.00014610985817853361

step: 90, loss: 0.0015730052255094051

step: 100, loss: 0.009988762438297272

step: 110, loss: 0.001359684974886477

step: 120, loss: 0.00011363014345988631

step: 130, loss: 0.000292594893835485

step: 140, loss: 0.0016539503121748567

step: 150, loss: 0.003128687385469675

step: 160, loss: 0.002491824561730027

step: 170, loss: 0.0022551838774234056

step: 180, loss: 5.407704520621337e-05

step: 190, loss: 0.0011334633454680443

step: 200, loss: 8.764705125940964e-05

step: 210, loss: 2.852586840162985e-05

step: 220, loss: 2.3863303795224056e-05

step: 230, loss: 0.025562459602952003

step: 240, loss: 0.0004268925404176116

step: 250, loss: 0.0001969145523617044

step: 260, loss: 0.0004929624265059829

step: 270, loss: 0.0007489097188226879

step: 280, loss: 0.014561534859240055

step: 290, loss: 0.00016972672892734408

step: 300, loss: 6.296703213592991e-05

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

step: 310, loss: 7.325013575609773e-05

step: 320, loss: 0.000178137983311899

step: 330, loss: 0.0002130098146153614

step: 340, loss: 0.011680756695568562

step: 350, loss: 3.690471203299239e-05

step: 360, loss: 0.10758765041828156

step: 370, loss: 0.0011110244086012244

step: 380, loss: 0.00024638487957417965

step: 390, loss: 0.00027408014284446836

step: 400, loss: 0.00012523954501375556

step: 410, loss: 0.0002354419557377696

step: 420, loss: 0.0015768549637869

step: 430, loss: 0.00012641798821277916

```

step: 440, loss: 0.026612892746925354
step: 450, loss: 0.007611579727381468
step: 460, loss: 0.0009166447562165558
step: 470, loss: 0.008434100076556206
step: 480, loss: 6.915586709510535e-05
step: 490, loss: 0.00016457363381050527
step: 500, loss: 0.00020248170767445117
step: 510, loss: 0.00016838000738061965
step: 520, loss: 0.0001372146507492289
step: 530, loss: 0.00010851269325939938
/usr/local/lib/python3.7/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=0.9550870760769936, f1=0.9468377635197067,
best_f1=0.946927374301676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007581817335449159
step: 10, loss: 0.009971125982701778
step: 20, loss: 0.0008147162152454257
step: 30, loss: 0.0007688504410907626
step: 40, loss: 0.00011653559340629727
step: 50, loss: 0.00679033575579524
step: 60, loss: 0.0003273698966950178
step: 70, loss: 0.00020115984079893678
step: 80, loss: 0.00033698725746944547
step: 90, loss: 5.891668115509674e-05
step: 100, loss: 6.804908480262384e-05
step: 110, loss: 0.0007507430855184793
step: 120, loss: 0.0011104738805443048
step: 130, loss: 7.701648428337649e-05
step: 140, loss: 0.000703820725902915
step: 150, loss: 0.004341394640505314
step: 160, loss: 0.0019840181339532137
step: 170, loss: 0.008511491119861603
step: 180, loss: 0.013269586488604546
step: 190, loss: 7.421277405228466e-05
step: 200, loss: 3.461625237832777e-05
step: 210, loss: 0.0005717383464798331
step: 220, loss: 9.094073175219819e-05
step: 230, loss: 0.0003954804560635239
step: 240, loss: 0.0006847947952337563
step: 250, loss: 0.0003770585171878338
step: 260, loss: 0.0015064652543514967
step: 270, loss: 0.00014079012908041477
step: 280, loss: 0.00011623138561844826
step: 290, loss: 9.215913451043889e-05

```

```

step: 300, loss: 7.632044435013086e-05
step: 310, loss: 0.00456457631662488
step: 320, loss: 3.172634751535952e-05
step: 330, loss: 0.006719905883073807
step: 340, loss: 6.864197348477319e-05
step: 350, loss: 5.16452309966553e-05
step: 360, loss: 7.637379167135805e-05
step: 370, loss: 6.617992767132819e-05
step: 380, loss: 0.057279981672763824
step: 390, loss: 0.0008399957441724837
step: 400, loss: 0.016016928479075432
step: 410, loss: 0.0001721649314276874
step: 420, loss: 8.218910807045177e-05
step: 430, loss: 0.0014320388436317444
step: 440, loss: 0.00012410955969244242
step: 450, loss: 0.00015288201393559575
step: 460, loss: 0.023382985964417458
step: 470, loss: 5.8641249779611826e-05
step: 480, loss: 0.0003923738549929112
step: 490, loss: 0.00011284795618848875
step: 500, loss: 0.03622188791632652
step: 510, loss: 0.03582530468702316
step: 520, loss: 0.0006029403302818537
step: 530, loss: 0.000292245764285326
/usr/local/lib/python3.7/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=0.9557933922754769, f1=0.9469767441860464,
best_f1=0.946927374301676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006643016822636127
step: 10, loss: 0.001246385625563562
step: 20, loss: 0.0005491220508702099
step: 30, loss: 0.015390283428132534
step: 40, loss: 6.991186091909185e-05
step: 50, loss: 0.0011798745254054666
step: 60, loss: 0.0002923196880146861
step: 70, loss: 0.00021629620459862053
step: 80, loss: 0.00015307942521758378
step: 90, loss: 0.0008070521289482713
step: 100, loss: 0.00012259242066647857
step: 110, loss: 0.0034772322978824377
step: 120, loss: 0.0004840211768168956
step: 130, loss: 0.0002486038429196924
step: 140, loss: 0.00011744885705411434
step: 150, loss: 0.00015175332373473793

```

```

step: 160, loss: 8.864323172019795e-05
step: 170, loss: 0.0004626896989066154
step: 180, loss: 0.00015097801224328578
step: 190, loss: 0.0015215970342978835
step: 200, loss: 0.002287720562890172
step: 210, loss: 0.00048242296907119453
step: 220, loss: 8.742068166611716e-05
step: 230, loss: 0.00017989928892347962
step: 240, loss: 0.0021704896353185177
step: 250, loss: 0.0013503357768058777
step: 260, loss: 6.319374369923025e-05
step: 270, loss: 4.5196495193522424e-05
step: 280, loss: 0.0002126301551470533
step: 290, loss: 0.002212636638432741
step: 300, loss: 0.0006394258234649897
step: 310, loss: 0.0478607714176178
step: 320, loss: 0.00018502688908483833
step: 330, loss: 9.463581227464601e-05
step: 340, loss: 0.00039654740248806775
step: 350, loss: 0.006387058179825544
step: 360, loss: 0.0007062930963002145
step: 370, loss: 0.00043147781980223954
step: 380, loss: 0.00021091822418384254
step: 390, loss: 0.0001507515844423324
step: 400, loss: 0.002281190361827612
step: 410, loss: 0.000540520588401705
step: 420, loss: 0.0003754254139494151
step: 430, loss: 5.592719026026316e-05
step: 440, loss: 0.0002568235795479268
step: 450, loss: 0.009319442324340343
step: 460, loss: 0.007585258223116398
step: 470, loss: 5.2655916078947484e-05
step: 480, loss: 0.0016844626516103745
step: 490, loss: 0.003733657067641616
step: 500, loss: 0.0013996309135109186
step: 510, loss: 0.0001167948212241754
step: 520, loss: 5.426984716905281e-05
step: 530, loss: 0.00013132067397236824
/usr/local/lib/python3.7/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=0.9568044588945658, f1=0.9499072356215214,
best_f1=0.946927374301676
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:23, 249.43it/s]
load_f1 = 0.9622025198320111
real_f1 = 0.9626517273576097
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.18it/s]

```