

CMedium_30_3_5_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 | 21.12 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 50.9 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 74.7 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 90.2 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 60.1 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 32.8 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 | 9.37 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
  | 10.4 MB 4.2 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 71.9 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
  | 753.4 MB 6.7 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
  | 56 kB 5.7 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
  | 884 kB 67.7 MB/s
Collecting jsonlines==1.2.0

```

```

    Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)
Collecting nltk==3.4.5
    Downloading nltk-3.4.5.zip (1.5 MB)
      |                  | 1.5 MB 58.2 MB/s
Requirement already satisfied: six>=1.5.0 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.15.0)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
Collecting thinc<7.4.0,>=7.3.0
    Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
      |                  | 2.2 MB 58.3 MB/s
Collecting plac<1.2.0,>=0.9.6
    Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting blis<0.5.0,>=0.4.0
    Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
      |                  | 3.7 MB 54.3 MB/s
Collecting srsly<1.1.0,>=0.1.0
    Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
      |                  | 184 kB 78.0 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: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Collecting catalogue<1.1.0,>=0.0.7
    Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting tokenizers==0.8.1.rc2
    Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)

```

```

| 3.0 MB 48.6 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 sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    | 880 kB 34.3 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: 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: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449922
sha256=832b1d196cef49e7939711f3acae7a156ed63f4c1980a28aa631aa74513df0ba
  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=716c0effde8b2d1f4cf48fe7355e0e9e37250cefc6b17ea77071d672d47e767b
Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
Attempting uninstall: numpy
  Found existing installation: numpy 1.21.6
  Uninstalling numpy-1.21.6:
    Successfully uninstalled numpy-1.21.6
Attempting uninstall: tqdm
  Found existing installation: tqdm 4.64.0
  Uninstalling tqdm-4.64.0:
    Successfully uninstalled tqdm-4.64.0
Attempting uninstall: srsly
  Found existing installation: srsly 2.4.4
  Uninstalling srsly-2.4.4:
    Successfully uninstalled srsly-2.4.4
Attempting uninstall: regex
  Found existing installation: regex 2022.6.2
  Uninstalling regex-2022.6.2:
    Successfully uninstalled regex-2022.6.2
Attempting uninstall: blis
  Found existing installation: blis 0.7.8
  Uninstalling blis-0.7.8:
    Successfully uninstalled blis-0.7.8
Attempting uninstall: thinc
  Found existing installation: thinc 8.1.0
  Uninstalling thinc-8.1.0:
    Successfully uninstalled thinc-8.1.0
Attempting uninstall: catalogue
  Found existing installation: catalogue 2.0.8
  Uninstalling catalogue-2.0.8:
    Successfully uninstalled catalogue-2.0.8
Attempting uninstall: torch
  Found existing installation: torch 1.12.0+cu113
  Uninstalling torch-1.12.0+cu113:
    Successfully uninstalled torch-1.12.0+cu113
Attempting uninstall: spacy
  Found existing installation: spacy 3.4.1
  Uninstalling spacy-3.4.1:
    Successfully uninstalled spacy-3.4.1
Attempting uninstall: nltk
  Found existing installation: nltk 3.7
  Uninstalling nltk-3.7:

```

```

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

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

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

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

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

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

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

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

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

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

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

```

```
[1]: True
```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10117, done.
remote: Counting objects: 100% (233/233), done.
remote: Compressing objects: 100% (165/165), done.
remote: Total 10117 (delta 119), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.97 MiB | 22.18 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-phnt_2lk
Created temporary directory: /tmp/pip-req-tracker-w38l4sjl
Initialized build tracking at /tmp/pip-req-tracker-w38l4sjl
Created build tracker: /tmp/pip-req-tracker-w38l4sjl
Entered build tracker: /tmp/pip-req-tracker-w38l4sjl
```

```

Created temporary directory: /tmp/pip-install-hgvcfdto
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-hwrkrgc4
  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-w38l4sjl'
  Running setup.py (path:/tmp/pip-req-build-hwrkrgc4/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-5qluhenf
  Running command python setup.py egg_info

torch.__version__ = 1.4.0


running egg_info
creating /tmp/pip-pip-egg-info-5qluhenf/apex.egg-info
writing /tmp/pip-pip-egg-info-5qluhenf/apex.egg-info/PKG-INFO
writing dependency_links to /tmp/pip-pip-egg-info-5qluhenf/apex.egg-
info/dependency_links.txt
writing top-level names to /tmp/pip-pip-egg-info-5qluhenf/apex.egg-
info/top_level.txt
writing manifest file '/tmp/pip-pip-egg-info-5qluhenf/apex.egg-
info/SOURCES.txt'
adding license file 'LICENSE'
writing manifest file '/tmp/pip-pip-egg-info-5qluhenf/apex.egg-
info/SOURCES.txt'
Source in /tmp/pip-req-build-hwrkrgc4 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-w38l4sjl'
Created temporary directory: /tmp/pip-unpack-jwepvgo6
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-r5hm9h1k
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-r5hm9h1k
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''''/tmp/pip-req-build-hwrkrgc4/setup.py'''';

```



```
__file__='''/tmp/pip-req-build-hwrkrgc4/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-r5hm9hlk
```

```
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-r5hm9h1k/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=77b0d3406ff2bcd255fd8d4b3cff90c91327a96c5d71b0f3542c58b719a097c5
Stored in directory: /tmp/pip-ephem-wheel-cache-
phnt_2lk/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-w38l4sjl'
/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 7.6 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 62.6 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 10.2 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting botocore==1.27.48
  Downloading botocore-1.27.48-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 56.0 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 56.9 MB/s
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.48->awscli) (2.8.2)
Collecting jmespath<2.0.0,>=0.7.1
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.48->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
  Attempting uninstall: urllib3
    Found existing installation: urllib3 1.24.3
    Uninstalling urllib3-1.24.3:
      Successfully uninstalled urllib3-1.24.3
  Attempting uninstall: rsa
    Found existing installation: rsa 4.9
    Uninstalling rsa-4.9:
      Successfully uninstalled rsa-4.9
  Attempting uninstall: docutils
    Found existing installation: docutils 0.17.1
    Uninstalling docutils-0.17.1:
      Successfully uninstalled docutils-0.17.1
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.
panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is
incompatible.
fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is
incompatible.
Successfully installed awscli-1.25.48 botocore-1.27.48 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```

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

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
    |                                     | 776.8 MB 18 kB/s
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages
(from torch==1.7.1) (1.19.2)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 1014, done.
remote: Counting objects: 100% (84/84), done.
remote: Compressing objects: 100% (74/74), done.
remote: Total 1014 (delta 27), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (1014/1014), 254.12 MiB | 27.00 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/datasEsErros/CMedium_30_3_5/configs.json /content/
↪ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm 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, 653kB/s]

Downloading: 100% 899k/899k [00:00<00:00, 28.0MB/s]

Downloading: 100% 456k/456k [00:00<00:00, 22.8MB/s]

Downloading: 100% 501M/501M [00:10<00:00, 49.8MB/s]

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

Defaults for this optimization level are:

enabled : True

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4370655417442322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2692307692307693,
best_f1=0.2692307692307693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4294983744621277
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2666666666666667,
best_f1=0.2692307692307693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38396021723747253
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 3: dev_f1=0.34210526315789475, f1=0.3111111111111111,
best_f1=0.3111111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25043460726737976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4, f1=0.4799999999999999, best_f1=0.4799999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26980870962142944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.49122807017543857, f1=0.6, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16603343188762665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.7222222222222223, f1=0.6857142857142857,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3379935324192047
/usr/local/lib/python3.7/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=0.41791044776119407, f1=0.5957446808510638,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28282415866851807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7878787878787878, f1=0.7058823529411764,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20394828915596008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7741935483870968, f1=0.7096774193548386,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19168208539485931
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.7096774193548386,

```



```

best_f1=0.7096774193548386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09319324791431427
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8387096774193549, f1=0.75, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03616790473461151
/usr/local/lib/python3.7/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=0.8235294117647058, f1=0.7058823529411764, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05354250967502594
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7272727272727273, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03842749819159508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8387096774193549, f1=0.75, best_f1=0.75
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04991879686713219
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8387096774193549, f1=0.75, best_f1=0.75
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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, 108586.53it/s]
/usr/local/lib/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 = 0.7058823529411764
real_f1 = 0.7027027027027025
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:19, 221.77it/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: 0.6005876064300537
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.4345407783985138
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: 0.5726630091667175
step: 30, loss: 0.2759717106819153
step: 40, loss: 0.39234980940818787
step: 50, loss: 0.5565707087516785
step: 60, loss: 0.32151076197624207
step: 70, loss: 0.10444100201129913
step: 80, loss: 0.1567116230726242
step: 90, loss: 0.11700677871704102
step: 100, loss: 0.2192908525466919
step: 110, loss: 0.12317076325416565
step: 120, loss: 0.13849472999572754
step: 130, loss: 0.015688356012105942
step: 140, loss: 0.011002880521118641
step: 150, loss: 0.015072227455675602
step: 160, loss: 0.017418980598449707
step: 170, loss: 0.021292302757501602
step: 180, loss: 0.17858392000198364
step: 190, loss: 0.0257926844060421
step: 200, loss: 0.007198851555585861
step: 210, loss: 0.1176840215921402
step: 220, loss: 0.11066537350416183
step: 230, loss: 0.0036469013430178165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9718785151856018, f1=0.9705215419501134,
best_f1=0.9705215419501134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021314110606908798
step: 10, loss: 0.210979163646698
step: 20, loss: 0.032166652381420135
step: 30, loss: 0.007537170313298702
step: 40, loss: 0.07984242588281631
step: 50, loss: 0.02074011228978634
step: 60, loss: 0.021502966061234474
step: 70, loss: 0.01209726370871067

```

```

step: 80, loss: 0.004089401103556156
step: 90, loss: 0.0014624291798099875
step: 100, loss: 0.007274358533322811
step: 110, loss: 0.00824993010610342
step: 120, loss: 0.031239449977874756
step: 130, loss: 0.004872123245149851
step: 140, loss: 0.00410580774769187
step: 150, loss: 0.05795073136687279
step: 160, loss: 0.003787001594901085
step: 170, loss: 0.0013388121733441949
step: 180, loss: 0.0032426209654659033
step: 190, loss: 0.020421907305717468
step: 200, loss: 0.007969401776790619
step: 210, loss: 0.029007911682128906
step: 220, loss: 0.0009602121426723897
step: 230, loss: 0.0008238042937591672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819413092550789, f1=0.9762174405436014,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004797401372343302
step: 10, loss: 0.018113939091563225
step: 20, loss: 0.0008949976181611419
step: 30, loss: 0.03685460612177849
step: 40, loss: 0.006290160119533539
step: 50, loss: 0.014760306105017662
step: 60, loss: 0.002940040547400713
step: 70, loss: 0.0019560011569410563
step: 80, loss: 0.0007664301083423197
step: 90, loss: 0.004268209915608168
step: 100, loss: 0.01636551320552826
step: 110, loss: 0.018563903868198395
step: 120, loss: 0.0007193775963969529
step: 130, loss: 0.002732613356783986
step: 140, loss: 0.0005390682490542531
step: 150, loss: 0.06360502541065216
step: 160, loss: 0.00397457554936409
step: 170, loss: 0.003246545558795333
step: 180, loss: 0.018105067312717438
step: 190, loss: 0.006986362859606743

```

```

step: 200, loss: 0.008408332243561745
step: 210, loss: 0.0038704685866832733
step: 220, loss: 0.00395547179505229
step: 230, loss: 0.016887042671442032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, f1=0.9774266365688488,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006014863029122353
step: 10, loss: 0.009329511784017086
step: 20, loss: 0.01084950566291809
step: 30, loss: 0.0006429193308576941
step: 40, loss: 0.0932057574391365
step: 50, loss: 0.016049103811383247
step: 60, loss: 0.0036980623845010996
step: 70, loss: 0.0004507332341745496
step: 80, loss: 0.0009915796108543873
step: 90, loss: 0.0029768485110253096
step: 100, loss: 0.0011398118222132325
step: 110, loss: 0.0008047238225117326
step: 120, loss: 0.001163144945167005
step: 130, loss: 0.0020306340884417295
step: 140, loss: 0.01600836031138897
step: 150, loss: 0.00043084603385068476
step: 160, loss: 0.00784617941826582
step: 170, loss: 0.003213489195331931
step: 180, loss: 0.16715411841869354
step: 190, loss: 0.012621724046766758
step: 200, loss: 0.006686868146061897
step: 210, loss: 0.0012244838289916515
step: 220, loss: 0.010524558834731579
step: 230, loss: 0.00223249732516706
/usr/local/lib/python3.7/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=0.9807909604519773, f1=0.9818594104308391,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.005019110161811113
step: 10, loss: 0.0015497268177568913
step: 20, loss: 0.0006851240759715438
step: 30, loss: 0.00045669975224882364
step: 40, loss: 0.010013842955231667
step: 50, loss: 0.0006710966699756682
step: 60, loss: 0.004164968151599169
step: 70, loss: 0.0029700424056500196
step: 80, loss: 0.03206334263086319
step: 90, loss: 0.04701048135757446
step: 100, loss: 0.0008758658659644425
step: 110, loss: 0.0017598087433725595
step: 120, loss: 0.00651391688734293
step: 130, loss: 0.001331530511379242
step: 140, loss: 0.028779273852705956
step: 150, loss: 0.09230829030275345
step: 160, loss: 0.004474223591387272
step: 170, loss: 0.006289544049650431
step: 180, loss: 0.007936153560876846
step: 190, loss: 0.008575700223445892
step: 200, loss: 0.0006032832898199558
step: 210, loss: 0.002926694927737117
step: 220, loss: 0.0027988243382424116
step: 230, loss: 0.004166203085333109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987598647125141, 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: 0.002251311670988798
step: 10, loss: 0.00565325329080224
step: 20, loss: 0.0013113563181832433
step: 30, loss: 0.0006370781920850277
step: 40, loss: 0.0003273081674706191
step: 50, loss: 0.0004846765077672899
step: 60, loss: 0.0016171283787116408
step: 70, loss: 0.0003597326285671443
step: 80, loss: 0.0014025007840245962
step: 90, loss: 0.0069318669848144054
step: 100, loss: 0.00046490353997796774
step: 110, loss: 0.006715166382491589

```

```

step: 120, loss: 0.0004134780610911548
step: 130, loss: 0.000460866023786366
step: 140, loss: 0.0005954676889814436
step: 150, loss: 0.00014420623483601958
step: 160, loss: 0.02133793756365776
step: 170, loss: 0.00038187854806892574
step: 180, loss: 0.00013791551464237273
step: 190, loss: 0.00029695374541915953
step: 200, loss: 0.001597303431481123
step: 210, loss: 0.001894557848572731
step: 220, loss: 0.004947966430336237
step: 230, loss: 0.0018305904231965542
/usr/local/lib/python3.7/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=0.9876819708846584,
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: 0.0025600362569093704
step: 10, loss: 0.0008557740366086364
step: 20, loss: 0.0012596484739333391
step: 30, loss: 0.0013585002161562443
step: 40, loss: 0.004612314980477095
step: 50, loss: 0.0011229225201532245
step: 60, loss: 0.00037918283487670124
step: 70, loss: 0.2610381543636322
step: 80, loss: 0.012396267615258694
step: 90, loss: 0.003896792884916067
step: 100, loss: 0.0007229336188174784
step: 110, loss: 0.0011276145232841372
step: 120, loss: 0.0026218260172754526
step: 130, loss: 0.005723252426832914
step: 140, loss: 0.0009960067691281438
step: 150, loss: 0.006469155661761761
step: 160, loss: 0.00241175782866776
step: 170, loss: 0.0010705966269597411
step: 180, loss: 0.0004863765207119286
step: 190, loss: 0.0002935347147285938
step: 200, loss: 0.0008458061493001878
step: 210, loss: 0.01103464886546135
step: 220, loss: 0.00037351137143559754
step: 230, loss: 0.0009353089844807982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```


UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9887640449438202, f1=0.9864864864864865,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008967758622020483
step: 10, loss: 0.003050091676414013
step: 20, loss: 0.0029388568364083767
step: 30, loss: 0.0019612316973507404
step: 40, loss: 0.0028780223801732063
step: 50, loss: 0.0007206428563222289
step: 60, loss: 0.0008670032257214189
step: 70, loss: 0.00042915227822959423
step: 80, loss: 0.0308036170899868
step: 90, loss: 0.0008557414403185248
step: 100, loss: 0.0011593851959332824
step: 110, loss: 0.0017056787619367242
step: 120, loss: 0.0011613320093601942
step: 130, loss: 0.002455356763675809
step: 140, loss: 0.0007610503816977143
step: 150, loss: 0.06920036673545837
step: 160, loss: 0.0004990280140191317
step: 170, loss: 0.09357504546642303
step: 180, loss: 0.0010613296180963516
step: 190, loss: 0.001373936072923243
step: 200, loss: 0.0045599378645420074
step: 210, loss: 0.0017897762591019273
step: 220, loss: 0.0005800689104944468
step: 230, loss: 0.0003949144738726318
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9876265466816648,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004747141501866281
step: 10, loss: 0.001101777724191546
step: 20, loss: 0.0021728717256337404
step: 30, loss: 0.00441161822527647
step: 40, loss: 0.0004743289027828723
step: 50, loss: 0.0007446764502674341
step: 60, loss: 0.0008930402109399438
step: 70, loss: 0.06636980921030045
```

```

step: 80, loss: 0.00016021048941183835
step: 90, loss: 0.03333020955324173
step: 100, loss: 0.0003302470431663096
step: 110, loss: 0.0004557677311822772
step: 120, loss: 0.07740059494972229
step: 130, loss: 0.0026429120916873217
step: 140, loss: 0.003210898954421282
step: 150, loss: 0.0008226664504036307
step: 160, loss: 0.029661642387509346
step: 170, loss: 0.0004341225721873343
step: 180, loss: 0.00020575267262756824
step: 190, loss: 0.00010558625945122913
step: 200, loss: 0.00017473996558692306
step: 210, loss: 0.003002662444487214
step: 220, loss: 0.00017252190446015447
step: 230, loss: 0.00010284086602041498
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9821029082774049,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011465993884485215
step: 10, loss: 0.00014861967065371573
step: 20, loss: 9.902405145112425e-05
step: 30, loss: 6.406661123037338e-05
step: 40, loss: 0.002797531196847558
step: 50, loss: 0.00030398895614780486
step: 60, loss: 0.0007370676612481475
step: 70, loss: 0.0005117832915857434
step: 80, loss: 6.718617078149691e-05
step: 90, loss: 9.967006917577237e-05
step: 100, loss: 6.493658293038607e-05
step: 110, loss: 0.0013986864360049367
step: 120, loss: 6.319220119621605e-05
step: 130, loss: 0.00037138917832635343
step: 140, loss: 0.0006586899980902672
step: 150, loss: 0.002052045427262783
step: 160, loss: 5.372720625018701e-05
step: 170, loss: 0.00012608521501533687
step: 180, loss: 0.00016960236825980246
step: 190, loss: 9.724592382553965e-05
step: 200, loss: 0.0001658116962062195
step: 210, loss: 0.00025064218789339066
step: 220, loss: 0.00013299273268785328
step: 230, loss: 7.298769196495414e-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 10: dev_f1=0.9898989898989898, f1=0.9832402234636871,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.256498116068542e-05
step: 10, loss: 0.00010610974277369678
step: 20, loss: 0.00011511362390592694
step: 30, loss: 6.699843652313575e-05
step: 40, loss: 3.04480672639329e-05
step: 50, loss: 4.5621625758940354e-05
step: 60, loss: 0.00010268612822983414
step: 70, loss: 5.732563295168802e-05
step: 80, loss: 0.002026147674769163
step: 90, loss: 0.10545475780963898
step: 100, loss: 0.00763375498354435
step: 110, loss: 0.01618030294775963
step: 120, loss: 0.00019410514505580068
step: 130, loss: 9.920438606059179e-05
step: 140, loss: 0.0001346975623164326
step: 150, loss: 0.00012475985568016768
step: 160, loss: 0.005691020283848047
step: 170, loss: 0.0004641428531613201
step: 180, loss: 0.0002887898008339107
step: 190, loss: 7.441818888764828e-05
step: 200, loss: 0.0021955121774226427
step: 210, loss: 0.004495653323829174
step: 220, loss: 0.0001569599553477019
step: 230, loss: 5.5826872994657606e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9899216125419933, f1=0.9854096520763187,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 8.292902930406854e-05
step: 10, loss: 4.8918962420430034e-05
step: 20, loss: 0.023660490289330482
step: 30, loss: 0.01659618318080902
step: 40, loss: 6.70664303470403e-05
step: 50, loss: 0.00011103779979748651
step: 60, loss: 6.451793160522357e-05
step: 70, loss: 0.00019883387722074986
step: 80, loss: 5.16153231728822e-05
step: 90, loss: 0.0005344322999008
step: 100, loss: 6.870077777421102e-05
step: 110, loss: 4.437848474481143e-05
step: 120, loss: 5.504641740117222e-05
step: 130, loss: 7.584707054775208e-05
step: 140, loss: 5.740107371821068e-05
step: 150, loss: 6.907341594342142e-05
step: 160, loss: 0.010878631845116615
step: 170, loss: 0.0013155206106603146
step: 180, loss: 4.2538493289612234e-05
step: 190, loss: 0.01315239630639553
step: 200, loss: 3.131993435090408e-05
step: 210, loss: 8.153261296683922e-05
step: 220, loss: 0.016986606642603874
step: 230, loss: 6.560638576047495e-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=0.9899216125419933, f1=0.9854423292273236,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.784575528698042e-05
step: 10, loss: 0.0003249151341151446
step: 20, loss: 6.226230470929295e-05
step: 30, loss: 2.6808929760591127e-05
step: 40, loss: 5.4164625908015296e-05
step: 50, loss: 5.886247890884988e-05
step: 60, loss: 4.207345409668051e-05
step: 70, loss: 2.936428973043803e-05
step: 80, loss: 2.3486491045332514e-05
step: 90, loss: 3.120469409623183e-05
step: 100, loss: 6.04140805080533e-05
step: 110, loss: 9.325100836576894e-05
step: 120, loss: 4.4327531213639304e-05
step: 130, loss: 6.001477231620811e-05
step: 140, loss: 4.285995237296447e-05
step: 150, loss: 1.9325570974615403e-05

```

```

step: 160, loss: 0.025308439508080482
step: 170, loss: 5.454795973491855e-05
step: 180, loss: 0.0012570908293128014
step: 190, loss: 5.323756340658292e-05
step: 200, loss: 2.005932583415415e-05
step: 210, loss: 5.486497320816852e-05
step: 220, loss: 4.807109144167043e-05
step: 230, loss: 0.00021399730758275837
/usr/local/lib/python3.7/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=0.9899216125419933, f1=0.9864864864864865,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.6258807085687295e-05
step: 10, loss: 2.9982558771735057e-05
step: 20, loss: 0.0012262965319678187
step: 30, loss: 7.212837226688862e-05
step: 40, loss: 3.451210795901716e-05
step: 50, loss: 3.858284981106408e-05
step: 60, loss: 4.3309933971613646e-05
step: 70, loss: 3.460914740571752e-05
step: 80, loss: 4.305157926864922e-05
step: 90, loss: 6.110136018833145e-05
step: 100, loss: 4.1618652176111937e-05
step: 110, loss: 4.993742913939059e-05
step: 120, loss: 1.558601070428267e-05
step: 130, loss: 5.3167244914220646e-05
step: 140, loss: 3.3637305023148656e-05
step: 150, loss: 4.06179933634121e-05
step: 160, loss: 3.45052030752413e-05
step: 170, loss: 3.9220802136696875e-05
step: 180, loss: 4.0215694752987474e-05
step: 190, loss: 3.7957379390718415e-05
step: 200, loss: 4.600756074069068e-05
step: 210, loss: 2.6232097297906876e-05
step: 220, loss: 6.39580175629817e-05
step: 230, loss: 2.092757677019108e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9899216125419933, f1=0.9887640449438202,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0014097033999860287
step: 10, loss: 3.769745308090933e-05
step: 20, loss: 6.927783397259191e-05
step: 30, loss: 3.6818473745370284e-05
step: 40, loss: 2.2924265067558736e-05
step: 50, loss: 3.361120616318658e-05
step: 60, loss: 0.025953946635127068
step: 70, loss: 2.914136894105468e-05
step: 80, loss: 3.21400621032808e-05
step: 90, loss: 2.669402238097973e-05
step: 100, loss: 2.547949770814739e-05
step: 110, loss: 4.7691628424217924e-05
step: 120, loss: 0.0226425901055336
step: 130, loss: 2.7230402338318527e-05
step: 140, loss: 0.01934669353067875
step: 150, loss: 3.641654257080518e-05
step: 160, loss: 2.7964095352217555e-05
step: 170, loss: 1.9750506908167154e-05
step: 180, loss: 3.431538789300248e-05
step: 190, loss: 9.872261580312625e-05
step: 200, loss: 0.0017664864426478744
step: 210, loss: 0.032858990132808685
step: 220, loss: 3.7000765587436035e-05
step: 230, loss: 4.208544487482868e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9899216125419933, f1=0.9887892376681614,
best_f1=0.9854096520763187
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, 210.16it/s]
load_f1 = 0.9853768278965129
real_f1 = 0.9842696629213483
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 194.17it/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: 0.6428455114364624
```

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: 0.436378538608551
```

```
step: 20, loss: 0.3920472264289856
```

```
step: 30, loss: 0.31671714782714844
```

```
step: 40, loss: 0.2833322584629059
```

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

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

```
step: 50, loss: 1.070077896118164
```

```
step: 60, loss: 0.3247353732585907
```

```
step: 70, loss: 0.2664162814617157
```

```
step: 80, loss: 0.16604173183441162
```

```
step: 90, loss: 0.17757917940616608
```

```
step: 100, loss: 0.2391546070575714
```



```
step: 110, loss: 0.1717437207698822
step: 120, loss: 0.29878103733062744
step: 130, loss: 0.0975213572382927
step: 140, loss: 0.23760360479354858
step: 150, loss: 0.2501985430717468
step: 160, loss: 0.19869188964366913
step: 170, loss: 0.13719572126865387
step: 180, loss: 0.13052107393741608
step: 190, loss: 0.10149785131216049
step: 200, loss: 0.059414226561784744
step: 210, loss: 0.11074522882699966
step: 220, loss: 0.11725971102714539
step: 230, loss: 0.1370001882314682
step: 240, loss: 0.049905840307474136
step: 250, loss: 0.053800422698259354
step: 260, loss: 0.15877531468868256
step: 270, loss: 0.32333290576934814
step: 280, loss: 0.0854954719543457
step: 290, loss: 0.1435670107603073
step: 300, loss: 0.05103469640016556
step: 310, loss: 0.12384134531021118
step: 320, loss: 0.10911592096090317
step: 330, loss: 0.18416261672973633
step: 340, loss: 0.3610546588897705
step: 350, loss: 0.1679777204990387
step: 360, loss: 0.07055708765983582
step: 370, loss: 0.036135826259851456
step: 380, loss: 0.22095686197280884
step: 390, loss: 0.025992706418037415
step: 400, loss: 0.05191895365715027
step: 410, loss: 0.2814394533634186
step: 420, loss: 0.0466008298099041
step: 430, loss: 0.019946126267313957
step: 440, loss: 0.01708020456135273
step: 450, loss: 0.04015304520726204
step: 460, loss: 0.057416342198848724
step: 470, loss: 0.027925455942749977
step: 480, loss: 0.18375176191329956
step: 490, loss: 0.014219951815903187
step: 500, loss: 0.11895200610160828
step: 510, loss: 0.06581354886293411
step: 520, loss: 0.08467630296945572
step: 530, loss: 0.05154057592153549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.940795559666975, f1=0.9317972350230416,
best_f1=0.9317972350230416
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09071556478738785
step: 10, loss: 0.048416778445243835
step: 20, loss: 0.06987757235765457
step: 30, loss: 0.013401794247329235
step: 40, loss: 0.04430977627635002
step: 50, loss: 0.06556609272956848
step: 60, loss: 0.12296409159898758
step: 70, loss: 0.027293719351291656
step: 80, loss: 0.030915703624486923
step: 90, loss: 0.006236235611140728
step: 100, loss: 0.10217363387346268
step: 110, loss: 0.008330713026225567
step: 120, loss: 0.07289460301399231
step: 130, loss: 0.0140118058770895
step: 140, loss: 0.11198582500219345
step: 150, loss: 0.04370497167110443
step: 160, loss: 0.07711318880319595
step: 170, loss: 0.08033488690853119
step: 180, loss: 0.03491684049367905
step: 190, loss: 0.016213424503803253
step: 200, loss: 0.1649269014596939
step: 210, loss: 0.015330142341554165
step: 220, loss: 0.0009397593094035983
step: 230, loss: 0.041741251945495605
step: 240, loss: 0.04152462258934975
step: 250, loss: 0.04931073635816574
step: 260, loss: 0.05295509099960327
step: 270, loss: 0.11096250265836716
step: 280, loss: 0.04834398627281189
step: 290, loss: 0.03520923852920532
step: 300, loss: 0.01413422916084528
step: 310, loss: 0.019924458116292953
step: 320, loss: 0.1726522445678711
step: 330, loss: 0.06882220506668091
step: 340, loss: 0.03129159286618233
step: 350, loss: 0.005053170491009951
step: 360, loss: 0.11508022248744965
step: 370, loss: 0.037617579102516174
step: 380, loss: 0.15042924880981445
step: 390, loss: 0.011171405203640461
step: 400, loss: 0.04719995707273483

```

```

step: 410, loss: 0.06794720888137817
step: 420, loss: 0.046285342425107956
step: 430, loss: 0.16357043385505676
step: 440, loss: 0.012844924814999104
step: 450, loss: 0.04039726406335831
step: 460, loss: 0.04013751447200775
step: 470, loss: 0.026865556836128235
step: 480, loss: 0.018977178260684013
step: 490, loss: 0.08422830700874329
step: 500, loss: 0.0137609438970685
step: 510, loss: 0.03554972633719444
step: 520, loss: 0.4246916174888611
step: 530, loss: 0.09603389352560043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9411764705882353, f1=0.9344490934449092,
best_f1=0.9344490934449092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11084365099668503
step: 10, loss: 0.022083908319473267
step: 20, loss: 0.004749061074107885
step: 30, loss: 0.046191927045583725
step: 40, loss: 0.033986590802669525
step: 50, loss: 0.11481466889381409
step: 60, loss: 0.04661191254854202
step: 70, loss: 0.005030884873121977
step: 80, loss: 0.007007255684584379
step: 90, loss: 0.020683934912085533
step: 100, loss: 0.04876783490180969
step: 110, loss: 0.0767252966761589
step: 120, loss: 0.11470083892345428
step: 130, loss: 0.037711694836616516
step: 140, loss: 0.029119715094566345
step: 150, loss: 0.06648741662502289
step: 160, loss: 0.031190382316708565
step: 170, loss: 0.012268709018826485
step: 180, loss: 0.025930359959602356
step: 190, loss: 0.004266597796231508
step: 200, loss: 0.00931665301322937
step: 210, loss: 0.03017275594174862
step: 220, loss: 0.10154324024915695

```

```

step: 230, loss: 0.02166687324643135
step: 240, loss: 0.08882357180118561
step: 250, loss: 0.2268189638853073
step: 260, loss: 0.10052593052387238
step: 270, loss: 0.0048355781473219395
step: 280, loss: 0.0017934638308361173
step: 290, loss: 0.05598932132124901
step: 300, loss: 0.11402438580989838
step: 310, loss: 0.08059557527303696
step: 320, loss: 0.05328860506415367
step: 330, loss: 0.00840666238218546
step: 340, loss: 0.00949153769761324
step: 350, loss: 0.08721388876438141
step: 360, loss: 0.009627953171730042
step: 370, loss: 0.022065885365009308
step: 380, loss: 0.03848652169108391
step: 390, loss: 0.00350208580493927
step: 400, loss: 0.13361996412277222
step: 410, loss: 0.038122981786727905
step: 420, loss: 0.004489355254918337
step: 430, loss: 0.018754683434963226
step: 440, loss: 0.17703042924404144
step: 450, loss: 0.008556765504181385
step: 460, loss: 0.08866633474826813
step: 470, loss: 0.03787568211555481
step: 480, loss: 0.09636642783880234
step: 490, loss: 0.008474914357066154
step: 500, loss: 0.050067637115716934
step: 510, loss: 0.06417733430862427
step: 520, loss: 0.004100969526916742
step: 530, loss: 0.018521996214985847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9544186046511628, f1=0.9456928838951311,
best_f1=0.9456928838951311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010784360580146313
step: 10, loss: 0.0010160357924178243
step: 20, loss: 0.0936414822936058
step: 30, loss: 0.12846630811691284
step: 40, loss: 0.011540671810507774

```

step: 50, loss: 0.02767351269721985
step: 60, loss: 0.0109174232929945
step: 70, loss: 0.024306273087859154
step: 80, loss: 0.05170964077115059
step: 90, loss: 0.03406869247555733
step: 100, loss: 0.000970583816524595
step: 110, loss: 0.01708579994738102
step: 120, loss: 0.003059742273762822
step: 130, loss: 0.004116202238947153
step: 140, loss: 0.041512053459882736
step: 150, loss: 0.01617429405450821
step: 160, loss: 0.0015984044875949621
step: 170, loss: 0.01903085596859455
step: 180, loss: 0.12501004338264465
step: 190, loss: 0.04543760418891907
step: 200, loss: 0.07252948731184006
step: 210, loss: 0.004415198229253292
step: 220, loss: 0.002898281440138817
step: 230, loss: 0.029959101229906082
step: 240, loss: 0.03182126581668854
step: 250, loss: 0.1424063891172409
step: 260, loss: 0.012250254862010479
step: 270, loss: 0.06713180243968964
step: 280, loss: 0.014357421547174454
step: 290, loss: 0.07527118176221848
step: 300, loss: 0.002183838514611125
step: 310, loss: 0.016745280474424362
step: 320, loss: 0.07723191380500793
step: 330, loss: 0.015360258519649506
step: 340, loss: 0.008497586473822594
step: 350, loss: 0.04838789254426956
step: 360, loss: 0.044763240963220596
step: 370, loss: 0.003142973408102989
step: 380, loss: 0.01898556388914585
step: 390, loss: 0.0008148064953275025
step: 400, loss: 0.010018294677138329
step: 410, loss: 0.06671394407749176
step: 420, loss: 0.021702965721488
step: 430, loss: 0.03172372654080391
step: 440, loss: 0.006550562102347612
step: 450, loss: 0.006286136340349913
step: 460, loss: 0.046895332634449005
step: 470, loss: 0.0016749667702242732
step: 480, loss: 0.002959619741886854
step: 490, loss: 0.0005609566578641534
step: 500, loss: 0.007173325400799513
step: 510, loss: 0.03979984298348427
step: 520, loss: 0.09126976132392883

```
step: 530, loss: 0.1458308845758438
/usr/local/lib/python3.7/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=0.9499072356215214, f1=0.9465437788018433,
best_f1=0.9456928838951311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017911663744598627
step: 10, loss: 0.006139229983091354
step: 20, loss: 0.07069671154022217
step: 30, loss: 0.002575530903413892
step: 40, loss: 0.0005753772566094995
step: 50, loss: 0.011764845810830593
step: 60, loss: 0.003315706504508853
step: 70, loss: 0.001595210051164031
step: 80, loss: 0.0021653228905051947
step: 90, loss: 0.027002889662981033
step: 100, loss: 0.024415137246251106
step: 110, loss: 0.001157106482423842
step: 120, loss: 0.23258891701698303
step: 130, loss: 0.006754620466381311
step: 140, loss: 0.008406680077314377
step: 150, loss: 0.004752673674374819
step: 160, loss: 0.03133028373122215
step: 170, loss: 0.016237430274486542
step: 180, loss: 0.003946714103221893
step: 190, loss: 0.0032966528087854385
step: 200, loss: 0.009161331690847874
step: 210, loss: 0.0007793824188411236
step: 220, loss: 0.003276695031672716
step: 230, loss: 0.003717076499015093
step: 240, loss: 0.009633459150791168
step: 250, loss: 0.05798187106847763
step: 260, loss: 0.0016420497559010983
step: 270, loss: 0.0021989960223436356
step: 280, loss: 0.003125265473499894
step: 290, loss: 0.03451366722583771
step: 300, loss: 0.048961713910102844
step: 310, loss: 0.05552603676915169
step: 320, loss: 0.14459136128425598
step: 330, loss: 0.02370275929570198
step: 340, loss: 0.009390830993652344
step: 350, loss: 0.0008715293370187283
step: 360, loss: 0.0006227961275726557
step: 370, loss: 0.0010357891442254186
step: 380, loss: 0.009563571773469448
```

```

step: 390, loss: 0.010115444660186768
step: 400, loss: 0.0322122797369957
step: 410, loss: 0.029726078733801842
step: 420, loss: 0.23802100121974945
step: 430, loss: 0.0131898894906044
step: 440, loss: 0.0005177437560632825
step: 450, loss: 0.048117607831954956
step: 460, loss: 0.009546143002808094
step: 470, loss: 0.020527135580778122
step: 480, loss: 0.023882631212472916
step: 490, loss: 0.058916255831718445
step: 500, loss: 0.028023086488246918
step: 510, loss: 0.0011265479261055589
step: 520, loss: 0.1378631293773651
step: 530, loss: 0.0040959990583360195
/usr/local/lib/python3.7/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=0.952513966480447, f1=0.9419475655430711,
best_f1=0.9456928838951311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004036587662994862
step: 10, loss: 0.0022519631311297417
step: 20, loss: 0.01252794824540615
step: 30, loss: 0.0016198130324482918
step: 40, loss: 0.19010503590106964
step: 50, loss: 0.0016875243745744228
step: 60, loss: 0.0043105692602694035
step: 70, loss: 0.11432459950447083
step: 80, loss: 0.0006280098459683359
step: 90, loss: 0.004539988934993744
step: 100, loss: 0.005470810923725367
step: 110, loss: 0.003898969618603587
step: 120, loss: 0.17937034368515015
step: 130, loss: 0.0017440203810110688
step: 140, loss: 0.0008307466632686555
step: 150, loss: 0.0010450895642861724
step: 160, loss: 0.03562411665916443
step: 170, loss: 0.001544514438137412
step: 180, loss: 0.0027398974634706974
step: 190, loss: 0.16481514275074005
step: 200, loss: 0.008811675943434238
step: 210, loss: 0.004349452909082174
step: 220, loss: 0.004049328155815601
step: 230, loss: 0.002960752695798874
step: 240, loss: 0.04174908250570297

```

```

step: 250, loss: 0.017001036554574966
step: 260, loss: 0.0015109538799151778
step: 270, loss: 0.0033993786200881004
step: 280, loss: 0.0023763002827763557
step: 290, loss: 0.0007855013245716691
step: 300, loss: 0.008879845961928368
step: 310, loss: 0.056657824665308
step: 320, loss: 0.00026740459725260735
step: 330, loss: 0.008953725919127464
step: 340, loss: 0.04154288023710251
step: 350, loss: 0.0028737178072333336
step: 360, loss: 0.07264851778745651
step: 370, loss: 0.0025664675049483776
step: 380, loss: 0.00015689338033553213
step: 390, loss: 0.004024897236377001
step: 400, loss: 0.10850993543863297
step: 410, loss: 0.005964592099189758
step: 420, loss: 0.007522300351411104
step: 430, loss: 0.001195669872686267
step: 440, loss: 0.00023670891823712736
step: 450, loss: 0.16730307042598724
step: 460, loss: 0.003007170045748353
step: 470, loss: 0.031240040436387062
step: 480, loss: 0.006832359358668327
step: 490, loss: 0.005226545035839081
step: 500, loss: 0.0283577349036932
step: 510, loss: 0.2188146412372589
step: 520, loss: 0.0005636180867440999
step: 530, loss: 0.050264038145542145
/usr/local/lib/python3.7/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=0.9518015910154421, f1=0.9503280224929709,
best_f1=0.9456928838951311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002421777928248048
step: 10, loss: 0.0019747517071664333
step: 20, loss: 0.003423137590289116
step: 30, loss: 0.018827401101589203
step: 40, loss: 0.001811276888474822
step: 50, loss: 0.003653570543974638
step: 60, loss: 0.005888179410248995
step: 70, loss: 0.00023148265609052032
step: 80, loss: 0.0011741736670956016
step: 90, loss: 0.0002181615709559992
step: 100, loss: 0.09282669425010681

```



```

step: 110, loss: 0.004032832570374012
step: 120, loss: 0.004019351210445166
step: 130, loss: 0.0013629210880026221
step: 140, loss: 0.0028727364260703325
step: 150, loss: 0.002458631992340088
step: 160, loss: 0.001399464439600706
step: 170, loss: 0.0014634976396337152
step: 180, loss: 0.013547569513320923
step: 190, loss: 0.019488830119371414
step: 200, loss: 0.000657523749396205
step: 210, loss: 0.001693468657322228
step: 220, loss: 0.009934922680258751
step: 230, loss: 0.0017379655037075281
step: 240, loss: 0.06260666251182556
step: 250, loss: 0.02238328941166401
step: 260, loss: 0.01334551814943552
step: 270, loss: 0.0006763431592844427
step: 280, loss: 0.002618331229314208
step: 290, loss: 0.0012962794862687588
step: 300, loss: 0.0004323706089053303
step: 310, loss: 0.001352431601844728
step: 320, loss: 0.0029666235204786062
step: 330, loss: 0.025249386206269264
step: 340, loss: 0.008331498131155968
step: 350, loss: 0.0009129612590186298
step: 360, loss: 0.022937975823879242
step: 370, loss: 0.052108749747276306
step: 380, loss: 0.003859141608700156
step: 390, loss: 0.0032447988633066416
step: 400, loss: 0.016184119507670403
step: 410, loss: 0.003970160149037838
step: 420, loss: 0.11359325051307678
step: 430, loss: 0.002545698080211878
step: 440, loss: 0.00380340707488358
step: 450, loss: 0.01002802886068821
step: 460, loss: 0.0018137098522856832
step: 470, loss: 0.26111477613449097
step: 480, loss: 0.006088997703045607
step: 490, loss: 0.0075179473496973515
step: 500, loss: 0.008029531687498093
step: 510, loss: 0.0019462838536128402
step: 520, loss: 0.0036613685078918934
step: 530, loss: 0.002211366780102253
/usr/local/lib/python3.7/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=0.9520260829063809, f1=0.9478098788443615,
best_f1=0.9456928838951311

```

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

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

```
step: 0, loss: 0.0017381382640451193  
step: 10, loss: 0.00890401005744934  
step: 20, loss: 0.002976402174681425  
step: 30, loss: 0.00040873041143640876  
step: 40, loss: 0.00021729108993895352  
step: 50, loss: 0.0016175889177247882  
step: 60, loss: 0.0005602026358246803  
step: 70, loss: 0.0011201928136870265  
step: 80, loss: 0.0033820911776274443  
step: 90, loss: 0.00027085162582807243  
step: 100, loss: 0.0003278615768067539  
step: 110, loss: 0.0003636517212726176  
step: 120, loss: 0.0008248809026554227  
step: 130, loss: 0.0023144155275076628  
step: 140, loss: 0.0006214894237928092  
step: 150, loss: 0.0013423544587567449  
step: 160, loss: 0.00010248873149976134  
step: 170, loss: 0.14241817593574524  
step: 180, loss: 0.00020864132966380566  
step: 190, loss: 0.0009483560570515692  
step: 200, loss: 0.0013836428988724947  
step: 210, loss: 0.09075827151536942  
step: 220, loss: 0.0013792234240099788  
step: 230, loss: 0.0024686388205736876  
step: 240, loss: 0.00044653660734184086  
step: 250, loss: 0.0016653335187584162  
step: 260, loss: 0.00011926217121072114  
step: 270, loss: 0.002888477174565196  
step: 280, loss: 0.00011448210716480389  
step: 290, loss: 0.00011296752927592024  
step: 300, loss: 5.6010692787822336e-05  
step: 310, loss: 0.0001470624702051282  
step: 320, loss: 0.005577407777309418  
step: 330, loss: 0.0003279618395026773  
step: 340, loss: 0.1317559778690338  
step: 350, loss: 0.0002288206887897104  
step: 360, loss: 0.035937946289777756  
step: 370, loss: 0.008243966847658157  
step: 380, loss: 0.00014235831622499973  
step: 390, loss: 0.004663634579628706  
step: 400, loss: 0.002200178802013397  
step: 410, loss: 0.0014313748106360435  
step: 420, loss: 0.004353879950940609  
step: 430, loss: 0.0034282496199011803  
step: 440, loss: 0.020583108067512512
```

```

step: 450, loss: 0.0008856214699335396
step: 460, loss: 0.004972873721271753
step: 470, loss: 0.036747150123119354
step: 480, loss: 0.04297645762562752
step: 490, loss: 0.0019603674300014973
step: 500, loss: 0.0009992329869419336
step: 510, loss: 0.006462186574935913
step: 520, loss: 0.0007404112257063389
step: 530, loss: 0.0009190631681121886
/usr/local/lib/python3.7/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=0.9489322191272052, f1=0.9438515081206497,
best_f1=0.9456928838951311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012569117825478315
step: 10, loss: 0.00041733362013474107
step: 20, loss: 0.00023386803513858467
step: 30, loss: 0.04134334996342659
step: 40, loss: 0.0004958867793902755
step: 50, loss: 0.017297988757491112
step: 60, loss: 0.004336588084697723
step: 70, loss: 0.019266342744231224
step: 80, loss: 0.007960719056427479
step: 90, loss: 0.0762588158249855
step: 100, loss: 0.0006524701020680368
step: 110, loss: 0.0021105215419083834
step: 120, loss: 0.0018832116620615125
step: 130, loss: 0.0013382574543356895
step: 140, loss: 0.0003512138209771365
step: 150, loss: 0.0018693411257117987
step: 160, loss: 0.0009081231546588242
step: 170, loss: 0.0016785371117293835
step: 180, loss: 0.006161737721413374
step: 190, loss: 8.482330304104835e-05
step: 200, loss: 0.00011797394108725712
step: 210, loss: 0.0004924405948258936
step: 220, loss: 0.004519609268754721
step: 230, loss: 0.0005320970085449517
step: 240, loss: 0.00030526123009622097
step: 250, loss: 0.0003463192260824144
step: 260, loss: 0.0005903275450691581
step: 270, loss: 0.00790527742356062
step: 280, loss: 0.0018090368248522282
step: 290, loss: 0.00016685047012288123
step: 300, loss: 0.0004949549329467118

```

```

step: 310, loss: 0.008904779329895973
step: 320, loss: 0.0001108868164010346
step: 330, loss: 0.0004949251306243241
step: 340, loss: 0.0039734793826937675
step: 350, loss: 0.03241670876741409
step: 360, loss: 0.0007696668617427349
step: 370, loss: 0.00036813667975366116
step: 380, loss: 0.009839949198067188
step: 390, loss: 5.406954369391315e-05
step: 400, loss: 0.06194952130317688
step: 410, loss: 0.001322534866631031
step: 420, loss: 0.000358390825567767
step: 430, loss: 0.0004928457201458514
step: 440, loss: 0.01099467370659113
step: 450, loss: 0.002859319793060422
step: 460, loss: 0.000141443480970338
step: 470, loss: 0.006587877869606018
step: 480, loss: 0.0020389927085489035
step: 490, loss: 0.0003114438150078058
step: 500, loss: 0.00043844617903232574
step: 510, loss: 0.0037657844368368387
step: 520, loss: 0.002873774617910385
step: 530, loss: 0.0020726658403873444
/usr/local/lib/python3.7/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=0.9498580889309367, f1=0.9451795841209829,
best_f1=0.9456928838951311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002512797713279724
step: 10, loss: 0.0050746118649840355
step: 20, loss: 0.0003205843095201999
step: 30, loss: 0.004407827276736498
step: 40, loss: 0.0011371553409844637
step: 50, loss: 0.0005017908406443894
step: 60, loss: 0.0011681924806907773
step: 70, loss: 0.00024754501646384597
step: 80, loss: 0.0001701819128356874
step: 90, loss: 0.01624842919409275
step: 100, loss: 0.0012162530329078436
step: 110, loss: 0.010027850978076458
step: 120, loss: 0.00024846848100423813
step: 130, loss: 0.002712627872824669
step: 140, loss: 0.010663512162864208
step: 150, loss: 0.00018832214118447155
step: 160, loss: 0.021931439638137817

```

```

step: 170, loss: 0.002099870704114437
step: 180, loss: 0.0002117332478519529
step: 190, loss: 0.09235210716724396
step: 200, loss: 0.015548406168818474
step: 210, loss: 0.004683646839112043
step: 220, loss: 0.002876052400097251
step: 230, loss: 8.226809586631134e-05
step: 240, loss: 0.00018444101442582905
step: 250, loss: 0.001254342612810433
step: 260, loss: 0.0016283350996673107
step: 270, loss: 0.00013155445049051195
step: 280, loss: 0.0237441286444664
step: 290, loss: 0.00013938495249021798
step: 300, loss: 0.0016226700972765684
step: 310, loss: 0.02719958871603012
step: 320, loss: 0.008181188255548477
step: 330, loss: 0.0010512351291254163
step: 340, loss: 0.00013139254588168114
step: 350, loss: 0.00034708884777501225
step: 360, loss: 0.00016036244051065296
step: 370, loss: 0.0160478837788105
step: 380, loss: 0.0009264542022719979
step: 390, loss: 0.00032934240880422294
step: 400, loss: 0.00026697281282395124
step: 410, loss: 0.00024327379651367664
step: 420, loss: 0.00014189760258886963
step: 430, loss: 0.0002864408306777477
step: 440, loss: 0.00015524108312092721
step: 450, loss: 0.02152165211737156
step: 460, loss: 0.00020606376347132027
step: 470, loss: 0.003779681632295251
step: 480, loss: 0.010588285513222218
step: 490, loss: 0.0012584467185661197
step: 500, loss: 0.000728511658962816
step: 510, loss: 0.00011006057320628315
step: 520, loss: 0.0019626093562692404
step: 530, loss: 0.01786009781062603
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9575757575757574, f1=0.9491841491841492,
best_f1=0.9491841491841492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005397719796746969
step: 10, loss: 8.88911381480284e-05
step: 20, loss: 0.0001277662959182635
step: 30, loss: 0.00013788229261990637
step: 40, loss: 0.0002384827093919739
step: 50, loss: 0.00012959481682628393
step: 60, loss: 0.0069680726155638695
step: 70, loss: 9.036614937940612e-05
step: 80, loss: 6.118091550888494e-05
step: 90, loss: 0.00036563791218213737
step: 100, loss: 0.000463494478026405
step: 110, loss: 0.0023047581780701876
step: 120, loss: 0.00048751302529126406
step: 130, loss: 0.00018887191254179925
step: 140, loss: 0.00016189724556170404
step: 150, loss: 0.0011886986903846264
step: 160, loss: 8.486188744427636e-05
step: 170, loss: 2.9640250431839377e-05
step: 180, loss: 8.72536184033379e-05
step: 190, loss: 0.00020939834939781576
step: 200, loss: 4.247929609846324e-05
step: 210, loss: 0.00022466899827122688
step: 220, loss: 0.003944115713238716
step: 230, loss: 0.0001268559426534921
step: 240, loss: 0.008065053261816502
step: 250, loss: 0.003390273079276085
step: 260, loss: 0.00626661442220211
step: 270, loss: 0.0025281759444624186
step: 280, loss: 0.005054634064435959
step: 290, loss: 0.00035906920675188303
step: 300, loss: 0.0006340932450257242
step: 310, loss: 0.0003470387600827962
step: 320, loss: 0.0005186771741136909
step: 330, loss: 3.8922113162698224e-05
step: 340, loss: 0.00020359450718387961
step: 350, loss: 4.5195592974778265e-05
step: 360, loss: 0.0004928760463371873
step: 370, loss: 0.0016920168418437243
step: 380, loss: 0.00021851247583981603
step: 390, loss: 0.0002257609012303874
step: 400, loss: 8.969460031948984e-05
step: 410, loss: 0.0004179116222076118
step: 420, loss: 6.0295118601061404e-05
step: 430, loss: 0.0011126261670142412
step: 440, loss: 0.00021294120233505964
step: 450, loss: 0.0006380650447681546
step: 460, loss: 0.0022664412390440702
```

```

step: 470, loss: 0.00016209669411182404
step: 480, loss: 0.0004901653155684471
step: 490, loss: 0.00508438889372349
step: 500, loss: 9.644406236475334e-05
step: 510, loss: 0.0011123422300443053
step: 520, loss: 0.0009632160072214901
step: 530, loss: 0.009253681637346745
/usr/local/lib/python3.7/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=0.9471733086190918, f1=0.9431192660550459,
best_f1=0.9491841491841492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000436954025644809
step: 10, loss: 0.0032310769893229008
step: 20, loss: 0.0014593459200114012
step: 30, loss: 0.00010166969150304794
step: 40, loss: 0.0002329240378458053
step: 50, loss: 0.00011604510655160993
step: 60, loss: 8.004371920833364e-05
step: 70, loss: 9.712615428725258e-05
step: 80, loss: 0.2545364499092102
step: 90, loss: 0.009676985442638397
step: 100, loss: 0.0007802650798112154
step: 110, loss: 0.00030074172536842525
step: 120, loss: 0.0009135388536378741
step: 130, loss: 0.0008159307762980461
step: 140, loss: 0.012406392954289913
step: 150, loss: 0.011240012012422085
step: 160, loss: 0.040532831102609634
step: 170, loss: 0.003386058611795306
step: 180, loss: 0.00010036725871032104
step: 190, loss: 0.00021220583585090935
step: 200, loss: 0.016563810408115387
step: 210, loss: 0.011333351023495197
step: 220, loss: 9.139113535638899e-05
step: 230, loss: 0.003144505899399519
step: 240, loss: 0.0024494901299476624
step: 250, loss: 9.850761853158474e-05
step: 260, loss: 5.2740000683115795e-05
step: 270, loss: 0.0064943330362439156
step: 280, loss: 0.00018585183715913445
step: 290, loss: 0.0002571379882283509
step: 300, loss: 0.00016570965817663819
step: 310, loss: 0.00010997473873430863
step: 320, loss: 0.004354106727987528

```

```

step: 330, loss: 0.011516304686665535
step: 340, loss: 0.0019849108066409826
step: 350, loss: 2.6614632588461973e-05
step: 360, loss: 5.60809530725237e-05
step: 370, loss: 7.180229295045137e-05
step: 380, loss: 4.806019569514319e-05
step: 390, loss: 0.0018885097233578563
step: 400, loss: 0.0032492359168827534
step: 410, loss: 0.0003626826801337302
step: 420, loss: 3.78075274056755e-05
step: 430, loss: 0.00021736552298534662
step: 440, loss: 0.00294170412234962
step: 450, loss: 0.009122745133936405
step: 460, loss: 0.0004376620636321604
step: 470, loss: 0.0033771393354982138
step: 480, loss: 0.014114348217844963
step: 490, loss: 0.00017638840654399246
step: 500, loss: 6.06595967838075e-05
step: 510, loss: 0.001766461762599647
step: 520, loss: 0.0025681492406874895
step: 530, loss: 0.000340884318575263
/usr/local/lib/python3.7/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=0.9536607970342911, f1=0.9492151431209603,
best_f1=0.9491841491841492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006127231754362583
step: 10, loss: 6.188720726640895e-05
step: 20, loss: 8.159514982253313e-05
step: 30, loss: 2.7859625333803706e-05
step: 40, loss: 6.083952757762745e-05
step: 50, loss: 0.002229972742497921
step: 60, loss: 9.655234316596761e-05
step: 70, loss: 7.623298733960837e-05
step: 80, loss: 8.824377437122166e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.0001349256344838068
step: 100, loss: 0.003087046556174755
step: 110, loss: 0.0022212283220142126
step: 120, loss: 4.120251105632633e-05
step: 130, loss: 2.9865699616493657e-05
step: 140, loss: 0.00029001623624935746
step: 150, loss: 0.0009241732186637819
step: 160, loss: 9.497790597379208e-05
step: 170, loss: 0.004301822278648615

```



```

step: 180, loss: 0.0001349820231553167
step: 190, loss: 7.781950262142345e-05
step: 200, loss: 4.611936310539022e-05
step: 210, loss: 7.899887714302167e-05
step: 220, loss: 4.2180545278824866e-05
step: 230, loss: 0.00029889147845096886
step: 240, loss: 5.69321600778494e-05
step: 250, loss: 7.13925328454934e-05
step: 260, loss: 5.875614078831859e-05
step: 270, loss: 0.08034560084342957
step: 280, loss: 5.665449134539813e-05
step: 290, loss: 0.00014566126628778875
step: 300, loss: 8.871866157278419e-05
step: 310, loss: 0.00019344172324053943
step: 320, loss: 0.006830545142292976
step: 330, loss: 0.000341722130542621
step: 340, loss: 0.0012076868442818522
step: 350, loss: 1.4196553820511326e-05
step: 360, loss: 0.06549926847219467
step: 370, loss: 0.00018709659343585372
step: 380, loss: 0.0003149554831907153
step: 390, loss: 0.0012655439786612988
step: 400, loss: 0.00013861416664440185
step: 410, loss: 0.005233116913586855
step: 420, loss: 6.583402864634991e-05
step: 430, loss: 0.00023693102411925793
step: 440, loss: 4.2085415770998225e-05
step: 450, loss: 0.0008323561050929129
step: 460, loss: 0.000180178860318847
step: 470, loss: 0.00012784967839252204
step: 480, loss: 3.5180331906303763e-05
step: 490, loss: 0.00010802598990267143
step: 500, loss: 5.477893864735961e-05
step: 510, loss: 0.00020858852076344192
step: 520, loss: 0.0005415737978182733
step: 530, loss: 0.004068127367645502
/usr/local/lib/python3.7/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=0.9496268656716418, f1=0.9462867818776274,
best_f1=0.9491841491841492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002926560409832746
step: 10, loss: 0.001163668348453939
step: 20, loss: 0.00010603827831801027
step: 30, loss: 0.004489606246352196

```

step: 40, loss: 5.849526860401966e-05
step: 50, loss: 0.024504993110895157
step: 60, loss: 4.477081893128343e-05
step: 70, loss: 4.594403435476124e-05
step: 80, loss: 9.650111314840615e-05
step: 90, loss: 3.9539405406685546e-05
step: 100, loss: 4.8260400944855064e-05
step: 110, loss: 0.00034840425360016525
step: 120, loss: 4.4364594941725954e-05
step: 130, loss: 6.124685023678467e-05
step: 140, loss: 0.00039792899042367935
step: 150, loss: 6.19704369455576e-05
step: 160, loss: 6.469463551184162e-05
step: 170, loss: 9.242711530532688e-05
step: 180, loss: 5.7517590903444216e-05
step: 190, loss: 7.23324847058393e-05
step: 200, loss: 6.7705666879192e-05
step: 210, loss: 0.024369604885578156
step: 220, loss: 4.499250644585118e-05
step: 230, loss: 4.846036608796567e-05
step: 240, loss: 0.00039054802618920803
step: 250, loss: 0.0008680908940732479
step: 260, loss: 0.00014785610255785286
step: 270, loss: 0.00030384177807718515
step: 280, loss: 0.00014271136024035513
step: 290, loss: 2.8335991373751312e-05
step: 300, loss: 5.121417416376062e-05
step: 310, loss: 0.00022462962078861892
step: 320, loss: 5.119963316246867e-05
step: 330, loss: 0.00012248128768987954
step: 340, loss: 0.001845114747993648
step: 350, loss: 6.963602208998054e-05
step: 360, loss: 4.290796277928166e-05
step: 370, loss: 0.004985298495739698
step: 380, loss: 0.0001574742782395333
step: 390, loss: 0.004008554387837648
step: 400, loss: 0.009587435983121395
step: 410, loss: 2.919603866757825e-05
step: 420, loss: 0.0004766028723679483
step: 430, loss: 0.00023445405531674623
step: 440, loss: 0.00011939999967580661
step: 450, loss: 6.358874816214666e-05
step: 460, loss: 0.005650816950947046
step: 470, loss: 0.000809514953289181
step: 480, loss: 7.707967597525567e-05
step: 490, loss: 0.0004800260649062693
step: 500, loss: 0.0008371055009774864
step: 510, loss: 0.0005712178535759449

```

step: 520, loss: 0.00012584308569785208
step: 530, loss: 0.0006434397655539215
/usr/local/lib/python3.7/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=0.9524253731343284, f1=0.9467289719626168,
best_f1=0.9491841491841492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013380103337112814
step: 10, loss: 0.000519688124768436
step: 20, loss: 0.00042838414083234966
step: 30, loss: 0.0016926922835409641
step: 40, loss: 0.00018183472275268286
step: 50, loss: 0.0022841196041554213
step: 60, loss: 0.0005679387832060456
step: 70, loss: 7.547229324700311e-05
step: 80, loss: 7.017006282694638e-05
step: 90, loss: 5.862990292371251e-05
step: 100, loss: 1.9086386600974947e-05
step: 110, loss: 0.0005223239422775805
step: 120, loss: 6.0112033679615706e-05
step: 130, loss: 0.00017556818784214556
step: 140, loss: 6.261104863369837e-05
step: 150, loss: 0.0002712543064262718
step: 160, loss: 7.718682172708213e-05
step: 170, loss: 5.0641225243452936e-05
step: 180, loss: 0.00011137066758237779
step: 190, loss: 0.0004933031741529703
step: 200, loss: 0.0008576521067880094
step: 210, loss: 0.00023773554130457342
step: 220, loss: 7.218424434540793e-05
step: 230, loss: 0.24631720781326294
step: 240, loss: 9.487134957453236e-05
step: 250, loss: 5.172564851818606e-05
step: 260, loss: 5.532243085326627e-05
step: 270, loss: 9.726315329317003e-05
step: 280, loss: 7.086002733558416e-05
step: 290, loss: 5.147802221472375e-05
step: 300, loss: 0.004609516356140375
step: 310, loss: 0.04303043335676193
step: 320, loss: 0.00011678897135425359
step: 330, loss: 0.00012994305870961398
step: 340, loss: 0.0005355658358894289
step: 350, loss: 0.0007328983629122376
step: 360, loss: 6.164415390230715e-05
step: 370, loss: 0.00010000359907280654

```

```

step: 380, loss: 7.216943049570546e-05
step: 390, loss: 9.936308197211474e-05
step: 400, loss: 0.01256515085697174
step: 410, loss: 0.00045664928620681167
step: 420, loss: 5.6530418078182265e-05
step: 430, loss: 3.9666447264607996e-05
step: 440, loss: 8.019357483135536e-05
step: 450, loss: 0.03059791401028633
step: 460, loss: 0.00019440740288700908
step: 470, loss: 5.6340995797654614e-05
step: 480, loss: 0.0014746689703315496
step: 490, loss: 0.0011419735383242369
step: 500, loss: 0.011261099949479103
step: 510, loss: 7.670358172617853e-05
step: 520, loss: 0.0017114623915404081
step: 530, loss: 4.0167920815292746e-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=0.9519813519813519, f1=0.9472701819878675,
best_f1=0.9491841491841492
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:20, 275.82it/s]
load_f1 = 0.9544186046511628
real_f1 = 0.9514925373134329
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, 207.29it/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 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: 0.48703861236572266
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4372633099555969
step: 20, loss: 0.4277274012565613
step: 30, loss: 0.38341236114501953
step: 40, loss: 0.28291016817092896
step: 50, loss: 0.39987537264823914
step: 60, loss: 0.48679688572883606
step: 70, loss: 0.33430930972099304
step: 80, loss: 0.372348427772522
step: 90, loss: 0.27226659655570984
step: 100, loss: 0.2314242571592331
step: 110, loss: 0.3398303687572479
step: 120, loss: 0.3918023109436035

```

```

step: 130, loss: 0.2656960189342499
step: 140, loss: 0.46980318427085876
step: 150, loss: 0.37369003891944885
step: 160, loss: 0.49410662055015564
step: 170, loss: 0.24525731801986694
step: 180, loss: 0.3771621584892273
step: 190, loss: 0.5957652926445007
step: 200, loss: 0.397040992975235
step: 210, loss: 0.5420620441436768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3757440745830536
step: 10, loss: 0.20241248607635498
step: 20, loss: 0.5166983008384705
step: 30, loss: 0.46990135312080383
step: 40, loss: 0.44552719593048096
step: 50, loss: 0.23045092821121216
step: 60, loss: 0.3198981285095215
step: 70, loss: 0.45119988918304443
step: 80, loss: 0.3118818700313568
step: 90, loss: 0.39277493953704834
step: 100, loss: 0.5293835997581482
step: 110, loss: 0.3757587969303131
step: 120, loss: 0.23983502388000488
step: 130, loss: 0.17310838401317596
step: 140, loss: 0.2345464825630188
step: 150, loss: 0.42417484521865845
step: 160, loss: 0.15653164684772491
step: 170, loss: 0.5923997759819031
step: 180, loss: 0.3194909393787384
step: 190, loss: 0.30947524309158325
step: 200, loss: 0.14850708842277527
step: 210, loss: 0.3077012598514557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2451642006635666
step: 10, loss: 0.2371755987405777
step: 20, loss: 0.4841457009315491
step: 30, loss: 0.251263827085495
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.4605080485343933
step: 50, loss: 0.48426976799964905
step: 60, loss: 0.4987736940383911
step: 70, loss: 0.1991957575082779
step: 80, loss: 0.46862974762916565
step: 90, loss: 0.23879146575927734
step: 100, loss: 0.389672189950943
step: 110, loss: 0.23956641554832458
step: 120, loss: 0.23186980187892914
step: 130, loss: 0.15442048013210297
step: 140, loss: 0.37767094373703003
step: 150, loss: 0.3237311840057373
step: 160, loss: 0.20295314490795135
step: 170, loss: 0.37545251846313477
step: 180, loss: 0.2391950637102127
step: 190, loss: 0.16166311502456665
step: 200, loss: 0.23304584622383118
step: 210, loss: 0.2582980692386627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3453706204891205
step: 10, loss: 0.30933719873428345
step: 20, loss: 0.2936190962791443
step: 30, loss: 0.2416699379682541
step: 40, loss: 0.23439329862594604
step: 50, loss: 0.24960024654865265
step: 60, loss: 0.49758973717689514
step: 70, loss: 0.24814923107624054
step: 80, loss: 0.24035976827144623
step: 90, loss: 0.4496185779571533
step: 100, loss: 0.38118618726730347
step: 110, loss: 0.5867581367492676
step: 120, loss: 0.3803413212299347
step: 130, loss: 0.6726879477500916

```



```

step: 140, loss: 0.5013930797576904
step: 150, loss: 0.38548433780670166
step: 160, loss: 0.31357815861701965
step: 170, loss: 0.17513759434223175
step: 180, loss: 0.08939892053604126
step: 190, loss: 0.1659162938594818
step: 200, loss: 0.30582523345947266
step: 210, loss: 0.38481783866882324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37944886088371277
step: 10, loss: 0.31658992171287537
step: 20, loss: 0.31824928522109985
step: 30, loss: 0.23246335983276367
step: 40, loss: 0.42316141724586487
step: 50, loss: 0.3787674307823181
step: 60, loss: 0.39026597142219543
step: 70, loss: 0.2510892152786255
step: 80, loss: 0.43217605352401733
step: 90, loss: 0.42500582337379456
step: 100, loss: 0.24510735273361206
step: 110, loss: 0.1665959507226944
step: 120, loss: 0.2404485046863556
step: 130, loss: 0.3110901415348053
step: 140, loss: 0.515785813331604
step: 150, loss: 0.31370842456817627
step: 160, loss: 0.2436876893043518
step: 170, loss: 0.37529587745666504
step: 180, loss: 0.24153423309326172
step: 190, loss: 0.5439451336860657
step: 200, loss: 0.36992964148521423
step: 210, loss: 0.24611423909664154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16056574881076813
step: 10, loss: 0.3828463852405548

```

```

step: 20, loss: 0.3182421326637268
step: 30, loss: 0.23413124680519104
step: 40, loss: 0.30842965841293335
step: 50, loss: 0.6150900721549988
step: 60, loss: 0.31653720140457153
step: 70, loss: 0.45154595375061035
step: 80, loss: 0.3142932057380676
step: 90, loss: 0.31131261587142944
step: 100, loss: 0.31508126854896545
step: 110, loss: 0.3779354393482208
step: 120, loss: 0.24615079164505005
step: 130, loss: 0.23990210890769958
step: 140, loss: 0.37550732493400574
step: 150, loss: 0.19073431193828583
step: 160, loss: 0.16595245897769928
step: 170, loss: 0.4510861933231354
step: 180, loss: 0.3920387327671051
step: 190, loss: 0.4506545066833496
step: 200, loss: 0.31556251645088196
step: 210, loss: 0.372784823179245
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45664992928504944
step: 10, loss: 0.3092542290687561
step: 20, loss: 0.4545627534389496
step: 30, loss: 0.24663172662258148
step: 40, loss: 0.25348997116088867
step: 50, loss: 0.3760402202606201
step: 60, loss: 0.31277936697006226
step: 70, loss: 0.2485278695821762
step: 80, loss: 0.5063970685005188
step: 90, loss: 0.37100091576576233
step: 100, loss: 0.4335322380065918
step: 110, loss: 0.3081360161304474
step: 120, loss: 0.3018527626991272
step: 130, loss: 0.4683201313018799
step: 140, loss: 0.31496527791023254
step: 150, loss: 0.31963032484054565
step: 160, loss: 0.5652273893356323
step: 170, loss: 0.6376721262931824
step: 180, loss: 0.25031715631484985
step: 190, loss: 0.2438494712114334

```

```

step: 200, loss: 0.31476324796676636
step: 210, loss: 0.3159103989601135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6082519292831421
step: 10, loss: 0.3121248185634613
step: 20, loss: 0.25338393449783325
step: 30, loss: 0.25562286376953125
step: 40, loss: 0.23968090116977692
step: 50, loss: 0.16452111303806305
step: 60, loss: 0.1640169769525528
step: 70, loss: 0.4543125331401825
step: 80, loss: 0.31236085295677185
step: 90, loss: 0.3737366199493408
step: 100, loss: 0.605002760887146
step: 110, loss: 0.384113609790802
step: 120, loss: 0.3101346492767334
step: 130, loss: 0.17221282422542572
step: 140, loss: 0.3124088943004608
step: 150, loss: 0.45411455631256104
step: 160, loss: 0.45082545280456543
step: 170, loss: 0.5348711609840393
step: 180, loss: 0.3143744468688965
step: 190, loss: 0.23824520409107208
step: 200, loss: 0.4617118239402771
step: 210, loss: 0.37894222140312195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5708286166191101
step: 10, loss: 0.2537992298603058
step: 20, loss: 0.3737708330154419
step: 30, loss: 0.17656971514225006
step: 40, loss: 0.16766107082366943
step: 50, loss: 0.4536532461643219
step: 60, loss: 0.17417222261428833
step: 70, loss: 0.3811720311641693

```

```

step: 80, loss: 0.3179512917995453
step: 90, loss: 0.3083569407463074
step: 100, loss: 0.4582410156726837
step: 110, loss: 0.5393180847167969
step: 120, loss: 0.37872856855392456
step: 130, loss: 0.24186736345291138
step: 140, loss: 0.5969129204750061
step: 150, loss: 0.11203697323799133
step: 160, loss: 0.3860814571380615
step: 170, loss: 0.3143436312675476
step: 180, loss: 0.4521755278110504
step: 190, loss: 0.3139299154281616
step: 200, loss: 0.3132971227169037
step: 210, loss: 0.31449824571609497
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3113134503364563
step: 10, loss: 0.3780207931995392
step: 20, loss: 0.25283247232437134
step: 30, loss: 0.1698472499847412
step: 40, loss: 0.3071991503238678
step: 50, loss: 0.5078248977661133
step: 60, loss: 0.24526795744895935
step: 70, loss: 0.37664496898651123
step: 80, loss: 0.16624920070171356
step: 90, loss: 0.3847285807132721
step: 100, loss: 0.45269060134887695
step: 110, loss: 0.1802794188261032
step: 120, loss: 0.45618486404418945
step: 130, loss: 0.24269329011440277
step: 140, loss: 0.3752972185611725
step: 150, loss: 0.24708817899227142
step: 160, loss: 0.3765205144882202
step: 170, loss: 0.2504688799381256
step: 180, loss: 0.3186076283454895
step: 190, loss: 0.24713164567947388
step: 200, loss: 0.6515439748764038
step: 210, loss: 0.24646729230880737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.18519984170953702, f1=0.18519984170953702,

```

```

best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30904334783554077
step: 10, loss: 0.3191092014312744
step: 20, loss: 0.38280296325683594
step: 30, loss: 0.16964685916900635
step: 40, loss: 0.44382983446121216
step: 50, loss: 0.43502628803253174
step: 60, loss: 0.4389890730381012
step: 70, loss: 0.17433233559131622
step: 80, loss: 0.5194947719573975
step: 90, loss: 0.3874870538711548
step: 100, loss: 0.5251038074493408
step: 110, loss: 0.44089335203170776
step: 120, loss: 0.4462621510028839
step: 130, loss: 0.23719938099384308
step: 140, loss: 0.30763083696365356
step: 150, loss: 0.3072744905948639
step: 160, loss: 0.1648751050233841
step: 170, loss: 0.31366047263145447
step: 180, loss: 0.24323531985282898
step: 190, loss: 0.4591929316520691
step: 200, loss: 0.18135273456573486
step: 210, loss: 0.3726663589477539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24092510342597961
step: 10, loss: 0.3092138171195984
step: 20, loss: 0.45850569009780884
step: 30, loss: 0.3118878901004791
step: 40, loss: 0.1734471619129181
step: 50, loss: 0.44855043292045593
step: 60, loss: 0.10776621848344803
step: 70, loss: 0.4430095851421356
step: 80, loss: 0.3134494125843048
step: 90, loss: 0.38666921854019165
step: 100, loss: 0.12769252061843872
step: 110, loss: 0.24606254696846008
step: 120, loss: 0.2502076327800751
step: 130, loss: 0.4443376660346985

```

```

step: 140, loss: 0.38425305485725403
step: 150, loss: 0.1021689772605896
step: 160, loss: 0.24689967930316925
step: 170, loss: 0.2410484105348587
step: 180, loss: 0.31431689858436584
step: 190, loss: 0.37501105666160583
step: 200, loss: 0.1730794459581375
step: 210, loss: 0.311300665140152
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31191906332969666
step: 10, loss: 0.24180461466312408
step: 20, loss: 0.3091887831687927
step: 30, loss: 0.24509698152542114
step: 40, loss: 0.24681290984153748
step: 50, loss: 0.24699999392032623
step: 60, loss: 0.5725386142730713
step: 70, loss: 0.43967315554618835
step: 80, loss: 0.37979400157928467
step: 90, loss: 0.24581508338451385
step: 100, loss: 0.38011476397514343
step: 110, loss: 0.23763605952262878
step: 120, loss: 0.3138800263404846
step: 130, loss: 0.3129939138889313
step: 140, loss: 0.16218361258506775
step: 150, loss: 0.31212422251701355
step: 160, loss: 0.5327454209327698
step: 170, loss: 0.23767292499542236
step: 180, loss: 0.24126584827899933
step: 190, loss: 0.2411515861749649
step: 200, loss: 0.37230947613716125
step: 210, loss: 0.38392174243927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24454770982265472
step: 10, loss: 0.38524767756462097

```

```

step: 20, loss: 0.5075430274009705
step: 30, loss: 0.2520592510700226
step: 40, loss: 0.3117053508758545
step: 50, loss: 0.24626126885414124
step: 60, loss: 0.3813238739967346
step: 70, loss: 0.2531658411026001
step: 80, loss: 0.308134526014328
step: 90, loss: 0.18042612075805664
step: 100, loss: 0.244638592004776
step: 110, loss: 0.45274582505226135
step: 120, loss: 0.24609923362731934
step: 130, loss: 0.31528493762016296
step: 140, loss: 0.30777910351753235
step: 150, loss: 0.3128507733345032
step: 160, loss: 0.1724320352077484
step: 170, loss: 0.45436733961105347
step: 180, loss: 0.24278387427330017
step: 190, loss: 0.38212040066719055
step: 200, loss: 0.16864942014217377
step: 210, loss: 0.4436681568622589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24112240970134735
step: 10, loss: 0.24367056787014008
step: 20, loss: 0.5118926167488098
step: 30, loss: 0.231472447514534
step: 40, loss: 0.3856126070022583
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.17336484789848328
step: 60, loss: 0.31686800718307495
step: 70, loss: 0.44870591163635254
step: 80, loss: 0.24088053405284882
step: 90, loss: 0.4568709135055542
step: 100, loss: 0.3105657994747162
step: 110, loss: 0.31512197852134705
step: 120, loss: 0.45711708068847656
step: 130, loss: 0.171123206615448
step: 140, loss: 0.316692054271698
step: 150, loss: 0.5225810408592224
step: 160, loss: 0.37596607208251953
step: 170, loss: 0.37474966049194336
step: 180, loss: 0.31099170446395874

```

```

step: 190, loss: 0.305527001619339
step: 200, loss: 0.3037177622318268
step: 210, loss: 0.4471181035041809
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 516.91it/s]
load_f1 = 0.18519984170953702
real_f1 = 0.18519984170953702
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:20, 215.80it/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: 0.46749117970466614
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4711432158946991
step: 20, loss: 0.25864213705062866
step: 30, loss: 0.3859328031539917
step: 40, loss: 0.23194637894630432
step: 50, loss: 0.3062077462673187
step: 60, loss: 0.44431716203689575
step: 70, loss: 0.46816691756248474
step: 80, loss: 0.19567257165908813
step: 90, loss: 0.31527912616729736
step: 100, loss: 0.43343082070350647
step: 110, loss: 0.2556871771812439
step: 120, loss: 0.33639729022979736
step: 130, loss: 0.3177908658981323
step: 140, loss: 0.1759427785873413
step: 150, loss: 0.3341194689273834
step: 160, loss: 0.23342496156692505
step: 170, loss: 0.3845243453979492
step: 180, loss: 0.15773622691631317
step: 190, loss: 0.15927042067050934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.3828509449958801
step: 10, loss: 0.31999513506889343
step: 20, loss: 0.660132110118866
step: 30, loss: 0.2632036507129669
step: 40, loss: 0.5645864605903625
step: 50, loss: 0.32047346234321594
step: 60, loss: 0.44959011673927307
step: 70, loss: 0.3074430227279663
step: 80, loss: 0.15015187859535217
step: 90, loss: 0.32117509841918945
step: 100, loss: 0.24748340249061584
step: 110, loss: 0.38001346588134766
step: 120, loss: 0.2435767948627472
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.4958072602748871
step: 140, loss: 0.2995646595954895
step: 150, loss: 0.28761813044548035
step: 160, loss: 0.23232366144657135
step: 170, loss: 0.2776481509208679
step: 180, loss: 0.15177883207798004
step: 190, loss: 0.20545470714569092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.40704500978473573, f1=0.4448051948051948,
best_f1=0.4448051948051948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3693639636039734
step: 10, loss: 0.27162814140319824
step: 20, loss: 0.3473513126373291
step: 30, loss: 0.2902079224586487
step: 40, loss: 0.11084762215614319
step: 50, loss: 0.17319047451019287
step: 60, loss: 0.16656973958015442
step: 70, loss: 0.19301128387451172

```

```

step: 80, loss: 0.15352784097194672
step: 90, loss: 0.13251237571239471
step: 100, loss: 0.21030589938163757
step: 110, loss: 0.4224573075771332
step: 120, loss: 0.2866271734237671
step: 130, loss: 0.04688937962055206
step: 140, loss: 0.0455731600522995
step: 150, loss: 0.3166724741458893
step: 160, loss: 0.13323499262332916
step: 170, loss: 0.22951585054397583
step: 180, loss: 0.1956261396408081
step: 190, loss: 0.010864981450140476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7934782608695652, f1=0.7859078590785908,
best_f1=0.7859078590785908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14519478380680084
step: 10, loss: 0.11885262280702591
step: 20, loss: 0.1601632982492447
step: 30, loss: 0.026522550731897354
step: 40, loss: 0.15313807129859924
step: 50, loss: 0.05650584027171135
step: 60, loss: 0.02916325442492962
step: 70, loss: 0.03717274218797684
step: 80, loss: 0.04149243235588074
step: 90, loss: 0.045339375734329224
step: 100, loss: 0.0752648189663887
step: 110, loss: 0.19584617018699646
step: 120, loss: 0.0851249247789383
step: 130, loss: 0.22812460362911224
step: 140, loss: 0.1447087973356247
step: 150, loss: 0.15685966610908508
step: 160, loss: 0.025841455906629562
step: 170, loss: 0.07957150787115097
step: 180, loss: 0.08842317759990692
step: 190, loss: 0.06457951664924622
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

epoch 4: dev_f1=0.8392370572207084, f1=0.8100558659217876,

best_f1=0.8100558659217876

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

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

step: 0, loss: 0.022151481360197067

step: 10, loss: 0.16518868505954742

step: 20, loss: 0.04434782266616821

step: 30, loss: 0.10920198261737823

step: 40, loss: 0.11940715461969376

step: 50, loss: 0.041834764182567596

step: 60, loss: 0.0034494802821427584

step: 70, loss: 0.07771698385477066

step: 80, loss: 0.03419042378664017

step: 90, loss: 0.047625284641981125

step: 100, loss: 0.0622994489967823

step: 110, loss: 0.1794556826353073

step: 120, loss: 0.014058029279112816

step: 130, loss: 0.22070378065109253

step: 140, loss: 0.018850967288017273

step: 150, loss: 0.0857955664396286

step: 160, loss: 0.03788745775818825

step: 170, loss: 0.02341519482433796

step: 180, loss: 0.020264923572540283

step: 190, loss: 0.04731239378452301

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

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

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

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=0.8556701030927835, f1=0.8177083333333333,

best_f1=0.8177083333333333

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

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

step: 0, loss: 0.15774069726467133

step: 10, loss: 0.058100562542676926

step: 20, loss: 0.01300124078989029

step: 30, loss: 0.06581227481365204

step: 40, loss: 0.013703505508601665

step: 50, loss: 0.022301677614450455

step: 60, loss: 0.020022790879011154

step: 70, loss: 0.01802227832376957

```

step: 80, loss: 0.06287568062543869
step: 90, loss: 0.07106364518404007
step: 100, loss: 0.14054393768310547
step: 110, loss: 0.005402375478297472
step: 120, loss: 0.007644682191312313
step: 130, loss: 0.008206469938158989
step: 140, loss: 0.006654699333012104
step: 150, loss: 0.005595089867711067
step: 160, loss: 0.04601164162158966
step: 170, loss: 0.12181323021650314
step: 180, loss: 0.05153147503733635
step: 190, loss: 0.0490235909819603
/usr/local/lib/python3.7/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=0.8333333333333334, f1=0.8099173553719008,
best_f1=0.8177083333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008545148186385632
step: 10, loss: 0.00647529074922204
step: 20, loss: 0.0007320214062929153
step: 30, loss: 0.10598912835121155
step: 40, loss: 0.13201899826526642
step: 50, loss: 0.000900412502232939
step: 60, loss: 0.058605264872312546
step: 70, loss: 0.0026217824779450893
step: 80, loss: 0.0011987551115453243
step: 90, loss: 0.0008326248498633504
step: 100, loss: 0.4419930577278137
step: 110, loss: 0.008875110186636448
step: 120, loss: 0.008125420659780502
step: 130, loss: 0.012012076564133167
step: 140, loss: 0.02315710298717022
step: 150, loss: 0.08094563335180283
step: 160, loss: 0.06643091142177582
step: 170, loss: 0.013111195527017117
step: 180, loss: 0.028076646849513054
step: 190, loss: 0.10239938646554947
/usr/local/lib/python3.7/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=0.8483290488431877, f1=0.8195876288659794,
best_f1=0.8177083333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.01644030772149563
step: 10, loss: 0.12139217555522919
step: 20, loss: 0.041170634329319
step: 30, loss: 0.14714182913303375
step: 40, loss: 0.05883647873997688
step: 50, loss: 0.05291755124926567
step: 60, loss: 0.0046848696656525135
step: 70, loss: 0.001380103058181703
step: 80, loss: 0.03199135512113571
step: 90, loss: 0.009517517872154713
step: 100, loss: 0.004738811869174242
step: 110, loss: 0.17641063034534454
step: 120, loss: 0.001972961239516735
step: 130, loss: 0.004408976528793573
step: 140, loss: 0.005972818937152624
step: 150, loss: 0.020381009206175804
step: 160, loss: 0.002167634665966034
step: 170, loss: 0.028828062117099762
step: 180, loss: 0.006580435670912266
step: 190, loss: 0.011119825765490532
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8571428571428572, f1=0.8195876288659794,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009036602452397346
step: 10, loss: 0.002360731363296509
step: 20, loss: 0.00265683070756495
step: 30, loss: 0.005194493569433689
step: 40, loss: 0.019757147878408432
step: 50, loss: 0.009029194712638855
step: 60, loss: 0.003594169393181801
step: 70, loss: 0.0006919695879332721
step: 80, loss: 0.20240825414657593
step: 90, loss: 0.0039252908900380135
step: 100, loss: 0.001857337774708867
step: 110, loss: 0.05620137229561806
step: 120, loss: 0.0033865920267999172
step: 130, loss: 0.03542625159025192
step: 140, loss: 0.15549565851688385
step: 150, loss: 0.00695383595302701

```

```

step: 160, loss: 0.004621313419193029
step: 170, loss: 0.14907410740852356
step: 180, loss: 0.009653238579630852
step: 190, loss: 0.0009720729431137443
/usr/local/lib/python3.7/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=0.8391959798994976, f1=0.8241469816272966,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038592941127717495
step: 10, loss: 0.005835734773427248
step: 20, loss: 0.07187049090862274
step: 30, loss: 0.1486852616071701
step: 40, loss: 0.006909720133990049
step: 50, loss: 0.028631199151277542
step: 60, loss: 0.0029131299816071987
step: 70, loss: 0.0011824237881228328
step: 80, loss: 0.003148280782625079
step: 90, loss: 0.0005948477773927152
step: 100, loss: 0.006860617082566023
step: 110, loss: 0.0009797642705962062
step: 120, loss: 0.002635700162500143
step: 130, loss: 0.011404616758227348
step: 140, loss: 0.0353541225194931
step: 150, loss: 0.007648400031030178
step: 160, loss: 0.0027500432915985584
step: 170, loss: 0.004081349354237318
step: 180, loss: 0.03755710646510124
step: 190, loss: 0.0018307029968127608
/usr/local/lib/python3.7/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=0.8350515463917526, f1=0.8144329896907216,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011144669726490974
step: 10, loss: 0.0007315626135095954
step: 20, loss: 0.004725002218037844
step: 30, loss: 0.04468480497598648
step: 40, loss: 0.0013239897089079022
step: 50, loss: 0.0013524852693080902
step: 60, loss: 0.0019253477221354842
step: 70, loss: 0.004928520414978266

```



```

step: 80, loss: 0.007080019451677799
step: 90, loss: 0.1534908264875412
step: 100, loss: 0.023590339347720146
step: 110, loss: 0.06556150317192078
step: 120, loss: 0.0012982224579900503
step: 130, loss: 0.009811780415475368
step: 140, loss: 0.001826282124966383
step: 150, loss: 0.0017780618509277701
step: 160, loss: 0.006892452482134104
step: 170, loss: 0.0010543158277869225
step: 180, loss: 0.01195554994046688
step: 190, loss: 0.03204677999019623
/usr/local/lib/python3.7/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=0.8491048593350383, f1=0.8443271767810027,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007693280465900898
step: 10, loss: 0.012127459980547428
step: 20, loss: 0.001607050304301083
step: 30, loss: 0.0058102114126086235
step: 40, loss: 0.003713061800226569
step: 50, loss: 0.017068596556782722
step: 60, loss: 0.0076842401176691055
step: 70, loss: 0.0025716214440762997
step: 80, loss: 0.003320883261039853
step: 90, loss: 0.0034530670382082462
step: 100, loss: 0.000645877153147012
step: 110, loss: 0.0034353199880570173
step: 120, loss: 0.0293634831905365
step: 130, loss: 0.0004950065049342811
step: 140, loss: 0.009122254326939583
step: 150, loss: 0.0059776464477181435
step: 160, loss: 0.0007011474808678031
step: 170, loss: 0.002379931742325425
step: 180, loss: 0.0002828651631716639
step: 190, loss: 0.0007204539142549038
/usr/local/lib/python3.7/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=0.8293963254593176, f1=0.8395721925133689,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.020473819226026535
step: 10, loss: 0.0005445372080430388
step: 20, loss: 0.00017846132686827332
step: 30, loss: 0.007527850568294525
step: 40, loss: 0.006762392818927765
step: 50, loss: 0.0011851171730086207
step: 60, loss: 0.024838166311383247
step: 70, loss: 0.02038346230983734
step: 80, loss: 0.00036977598210796714
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.10294406861066818
step: 100, loss: 0.01967497169971466
step: 110, loss: 0.02867473103106022
step: 120, loss: 0.0016936148749664426
step: 130, loss: 0.026112448424100876
step: 140, loss: 0.003592359833419323
step: 150, loss: 0.012214752845466137
step: 160, loss: 0.002401254838332534
step: 170, loss: 0.0009397217072546482
step: 180, loss: 0.0008684486383572221
step: 190, loss: 0.024700623005628586
/usr/local/lib/python3.7/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=0.8288770053475937, f1=0.8432432432432432,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004586989991366863
step: 10, loss: 0.006616570521146059
step: 20, loss: 0.0030095167458057404
step: 30, loss: 0.0010853464482352138
step: 40, loss: 0.0008637820719741285
step: 50, loss: 0.0228824932128191
step: 60, loss: 0.0038400008343160152
step: 70, loss: 0.0014139995910227299
step: 80, loss: 0.0008298391476273537
step: 90, loss: 0.0012053201207891107
step: 100, loss: 0.0010097200283780694
step: 110, loss: 0.25391143560409546
step: 120, loss: 0.002590768039226532
step: 130, loss: 0.0005509023903869092
step: 140, loss: 0.0011075454531237483
step: 150, loss: 0.004891573917120695
step: 160, loss: 0.0018176232697442174
step: 170, loss: 0.007557322271168232
step: 180, loss: 0.00396156869828701

```

```

step: 190, loss: 0.0020917938090860844
/usr/local/lib/python3.7/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=0.8232189973614776, f1=0.8409703504043127,
best_f1=0.8195876288659794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006534077692776918
step: 10, loss: 0.0009476529667153955
step: 20, loss: 0.0555666983127594
step: 30, loss: 0.0009252357413060963
step: 40, loss: 0.033563461154699326
step: 50, loss: 0.0007183132693171501
step: 60, loss: 0.007287212647497654
step: 70, loss: 0.0007314825197681785
step: 80, loss: 0.005439380649477243
step: 90, loss: 0.015212769620120525
step: 100, loss: 0.0012133773416280746
step: 110, loss: 0.02192508801817894
step: 120, loss: 0.0007072443258948624
step: 130, loss: 0.0008491299231536686
step: 140, loss: 0.0020823220256716013
step: 150, loss: 0.00040123151848092675
step: 160, loss: 0.0842042863368988
step: 170, loss: 0.05888392776250839
step: 180, loss: 0.0006296406500041485
step: 190, loss: 0.018684759736061096
/usr/local/lib/python3.7/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=0.828125, f1=0.8457446808510638, best_f1=0.8195876288659794
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:08, 236.72it/s]
load_f1 = 0.8179419525065963
real_f1 = 0.8188976377952756
733it [00:00, 3208.84it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 135.66it/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: 0.47827619314193726
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5314756631851196
step: 20, loss: 0.3522363603115082
step: 30, loss: 0.4422685503959656
step: 40, loss: 0.5095581412315369
step: 50, loss: 0.3370139002799988
step: 60, loss: 0.5604312419891357
step: 70, loss: 0.3249945342540741
step: 80, loss: 0.2282475382089615
step: 90, loss: 0.1921595335006714
step: 100, loss: 0.15583176910877228
step: 110, loss: 0.4070395231246948
step: 120, loss: 0.30008015036582947
step: 130, loss: 0.33538150787353516
step: 140, loss: 0.38045117259025574
step: 150, loss: 0.3034323453903198
step: 160, loss: 0.401550829410553
step: 170, loss: 0.33673977851867676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.31894782185554504
step: 10, loss: 0.48923298716545105
step: 20, loss: 0.35953840613365173
step: 30, loss: 0.32935404777526855
step: 40, loss: 0.08377423882484436
step: 50, loss: 0.42443224787712097
step: 60, loss: 0.1640920639038086
step: 70, loss: 0.4931207001209259
step: 80, loss: 0.24362021684646606
step: 90, loss: 0.2443619966506958
step: 100, loss: 0.5143795609474182
step: 110, loss: 0.2664680778980255
step: 120, loss: 0.25128597021102905
step: 130, loss: 0.5853283405303955
step: 140, loss: 0.5161676406860352
step: 150, loss: 0.41437309980392456

```

```

step: 160, loss: 0.4344285726547241
step: 170, loss: 0.39917492866516113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20027434842249658, f1=0.23107569721115537,
best_f1=0.23107569721115537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5806066393852234
step: 10, loss: 0.3254907429218292
step: 20, loss: 0.2275172770023346
step: 30, loss: 0.2550395429134369
step: 40, loss: 0.38357365131378174
step: 50, loss: 0.5857028365135193
step: 60, loss: 0.3027248978614807
step: 70, loss: 0.24366852641105652
step: 80, loss: 0.36870765686035156
step: 90, loss: 0.5757909417152405
step: 100, loss: 0.28955477476119995
step: 110, loss: 0.13827279210090637
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.4503336548805237
step: 130, loss: 0.4554498493671417
step: 140, loss: 0.27873310446739197
step: 150, loss: 0.150865375995636
step: 160, loss: 0.17235535383224487
step: 170, loss: 0.308504581451416
/usr/local/lib/python3.7/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=0.23107569721115537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3695036470890045
step: 10, loss: 0.5678054690361023
step: 20, loss: 0.22469036281108856
step: 30, loss: 0.452867716550827
step: 40, loss: 0.24225199222564697
step: 50, loss: 0.3449922502040863
step: 60, loss: 0.5749025940895081

```

```

step: 70, loss: 0.26910117268562317
step: 80, loss: 0.620756208896637
step: 90, loss: 0.3567522466182709
step: 100, loss: 0.39666131138801575
step: 110, loss: 0.3990367650985718
step: 120, loss: 0.5104129314422607
step: 130, loss: 0.23457255959510803
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.13119928538799286
step: 150, loss: 0.6680266261100769
step: 160, loss: 0.13574959337711334
step: 170, loss: 0.16745103895664215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6098901098901098, f1=0.5934065934065935,
best_f1=0.5934065934065935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12332476675510406
step: 10, loss: 0.31126871705055237
step: 20, loss: 0.33167994022369385
step: 30, loss: 0.3390037715435028
step: 40, loss: 0.25688889622688293
step: 50, loss: 0.2371821105480194
step: 60, loss: 0.3088320791721344
step: 70, loss: 0.18402285873889923
step: 80, loss: 0.35785990953445435
step: 90, loss: 0.23837006092071533
step: 100, loss: 0.13889193534851074
step: 110, loss: 0.26145175099372864
step: 120, loss: 0.11090783774852753
step: 130, loss: 0.15593573451042175
step: 140, loss: 0.13449867069721222
step: 150, loss: 0.30357998609542847
step: 160, loss: 0.21803201735019684
step: 170, loss: 0.11412443220615387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6571428571428571, f1=0.6652892561983471,
best_f1=0.6652892561983471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21608099341392517
step: 10, loss: 0.1635228544473648
step: 20, loss: 0.21922023594379425
step: 30, loss: 0.1209331676363945
step: 40, loss: 0.1160367876291275
step: 50, loss: 0.0908762663602829
step: 60, loss: 0.0990990698337555
step: 70, loss: 0.112077996134758
step: 80, loss: 0.10291191935539246
step: 90, loss: 0.14791667461395264
step: 100, loss: 0.104973703622818
step: 110, loss: 0.17501668632030487
step: 120, loss: 0.19500966370105743
step: 130, loss: 0.2612784504890442
step: 140, loss: 0.2406693398952484
step: 150, loss: 0.18437650799751282
step: 160, loss: 0.15403631329536438
step: 170, loss: 0.18676403164863586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7180722891566266, f1=0.6812652068126521,
best_f1=0.6812652068126521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.089011549949646
step: 10, loss: 0.21340513229370117
step: 20, loss: 0.3688625395298004
step: 30, loss: 0.12701065838336945
step: 40, loss: 0.03616338595747948
step: 50, loss: 0.0340816006064415
step: 60, loss: 0.18630240857601166
step: 70, loss: 0.15743057429790497
step: 80, loss: 0.08454042673110962
step: 90, loss: 0.08800237625837326
step: 100, loss: 0.04327787831425667
step: 110, loss: 0.11167503893375397

```

```

step: 120, loss: 0.11629495024681091
step: 130, loss: 0.1404402107000351
step: 140, loss: 0.06374004483222961
step: 150, loss: 0.13207173347473145
step: 160, loss: 0.06712555140256882
step: 170, loss: 0.10185717791318893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7647058823529412, f1=0.7553444180522566,
best_f1=0.7553444180522566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12569178640842438
step: 10, loss: 0.0869804248213768
step: 20, loss: 0.18401655554771423
step: 30, loss: 0.04145967587828636
step: 40, loss: 0.10320966690778732
step: 50, loss: 0.055644772946834564
step: 60, loss: 0.06129584461450577
step: 70, loss: 0.07895231992006302
step: 80, loss: 0.02334880270063877
step: 90, loss: 0.1991572380065918
step: 100, loss: 0.03192467242479324
step: 110, loss: 0.12435464560985565
step: 120, loss: 0.19227324426174164
step: 130, loss: 0.06494805216789246
step: 140, loss: 0.14227242767810822
step: 150, loss: 0.04171639308333397
step: 160, loss: 0.035567641258239746
step: 170, loss: 0.10224423557519913
/usr/local/lib/python3.7/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=0.7634660421545668, f1=0.7727272727272726,
best_f1=0.7553444180522566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022838084027171135
step: 10, loss: 0.1939198225736618
step: 20, loss: 0.04374432936310768
step: 30, loss: 0.11958733946084976

```

```

step: 40, loss: 0.2073901742696762
step: 50, loss: 0.028205858543515205
step: 60, loss: 0.11524546146392822
step: 70, loss: 0.17850278317928314
step: 80, loss: 0.022625582292675972
step: 90, loss: 0.08268392831087112
step: 100, loss: 0.11156272888183594
step: 110, loss: 0.13765934109687805
step: 120, loss: 0.2609314024448395
step: 130, loss: 0.1968134045600891
step: 140, loss: 0.06684735417366028
step: 150, loss: 0.2897653579711914
step: 160, loss: 0.14016152918338776
step: 170, loss: 0.16680867969989777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7751196172248804, f1=0.787037037037037,
best_f1=0.787037037037037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04746940732002258
step: 10, loss: 0.019333142787218094
step: 20, loss: 0.21901434659957886
step: 30, loss: 0.09393689781427383
step: 40, loss: 0.12524546682834625
step: 50, loss: 0.11747822910547256
step: 60, loss: 0.29316210746765137
step: 70, loss: 0.07393492013216019
step: 80, loss: 0.05287580192089081
step: 90, loss: 0.010139680467545986
step: 100, loss: 0.05798114836215973
step: 110, loss: 0.01110400352627039
step: 120, loss: 0.018172651529312134
step: 130, loss: 0.052939288318157196
step: 140, loss: 0.015682391822338104
step: 150, loss: 0.08170858770608902
step: 160, loss: 0.004297483712434769
step: 170, loss: 0.023834850639104843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.7792207792207791, f1=0.7928388746803069,
best_f1=0.7928388746803069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14287957549095154
step: 10, loss: 0.0061471667140722275
step: 20, loss: 0.022377341985702515
step: 30, loss: 0.09065014868974686
step: 40, loss: 0.00330843566916883
step: 50, loss: 0.07896408438682556
step: 60, loss: 0.14415177702903748
step: 70, loss: 0.05130711942911148
step: 80, loss: 0.011442191898822784
step: 90, loss: 0.02443922869861126
step: 100, loss: 0.2521412670612335
step: 110, loss: 0.024996599182486534
step: 120, loss: 0.012422353960573673
step: 130, loss: 0.0245506688952446
step: 140, loss: 0.3253364861011505
step: 150, loss: 0.006720529403537512
step: 160, loss: 0.07517996430397034
step: 170, loss: 0.1262059211730957
/usr/local/lib/python3.7/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=0.7755102040816327, f1=0.8068459657701711,
best_f1=0.7928388746803069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017930487170815468
step: 10, loss: 0.06707582622766495
step: 20, loss: 0.00553295249119401
step: 30, loss: 0.17692570388317108
step: 40, loss: 0.013212349265813828
step: 50, loss: 0.050483133643865585
step: 60, loss: 0.2547139823436737
step: 70, loss: 0.015907229855656624
step: 80, loss: 0.021158814430236816
step: 90, loss: 0.07624957710504532
step: 100, loss: 0.016965191811323166
step: 110, loss: 0.056681692600250244
step: 120, loss: 0.03226936608552933
step: 130, loss: 0.20409893989562988
```

```

step: 140, loss: 0.018588732928037643
step: 150, loss: 0.008948949165642262
step: 160, loss: 0.09137038141489029
step: 170, loss: 0.20010028779506683
/usr/local/lib/python3.7/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=0.765375854214123, f1=0.7822222222222223,
best_f1=0.7928388746803069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011117218062281609
step: 10, loss: 0.0038813319988548756
step: 20, loss: 0.004559213761240244
step: 30, loss: 0.034104958176612854
step: 40, loss: 0.017574692144989967
step: 50, loss: 0.11353220045566559
step: 60, loss: 0.029691345989704132
step: 70, loss: 0.1609361618757248
step: 80, loss: 0.005618160590529442
step: 90, loss: 0.02807820960879326
step: 100, loss: 0.02173713967204094
step: 110, loss: 0.025663206353783607
step: 120, loss: 0.011625200510025024
step: 130, loss: 0.019836759194731712
step: 140, loss: 0.06486961245536804
step: 150, loss: 0.014872798696160316
step: 160, loss: 0.0222477987408638
step: 170, loss: 0.005383268464356661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7815533980582524, f1=0.8085106382978723,
best_f1=0.8085106382978723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04149670526385307
step: 10, loss: 0.012806976214051247
step: 20, loss: 0.07675842195749283
step: 30, loss: 0.03661806508898735
step: 40, loss: 0.027883507311344147
step: 50, loss: 0.007277539931237698

```

```

step: 60, loss: 0.030208101496100426
step: 70, loss: 0.02536429464817047
step: 80, loss: 0.01311145443469286
step: 90, loss: 0.004774438217282295
step: 100, loss: 0.1235990822315216
step: 110, loss: 0.05728035792708397
step: 120, loss: 0.06273464858531952
step: 130, loss: 0.17152981460094452
step: 140, loss: 0.07673206180334091
step: 150, loss: 0.025732586160302162
step: 160, loss: 0.06582163274288177
step: 170, loss: 0.047611918300390244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.792079207920792, f1=0.8132387706855791,
best_f1=0.8132387706855791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003140874905511737
step: 10, loss: 0.01416783407330513
step: 20, loss: 0.0032979988027364016
step: 30, loss: 0.04981815069913864
step: 40, loss: 0.007203984074294567
step: 50, loss: 0.02398689277470112
step: 60, loss: 0.009063017554581165
step: 70, loss: 0.1788255125284195
step: 80, loss: 0.004884145222604275
step: 90, loss: 0.006311760749667883
step: 100, loss: 0.06453820317983627
step: 110, loss: 0.012452853843569756
step: 120, loss: 0.2151188850402832
step: 130, loss: 0.18229109048843384
step: 140, loss: 0.04466485232114792
step: 150, loss: 0.012777073308825493
step: 160, loss: 0.01050348486751318
step: 170, loss: 0.008974668569862843
/usr/local/lib/python3.7/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=0.7869674185463658, f1=0.8123515439429927,
best_f1=0.8132387706855791
mv: cannot stat '*_dev.pt': No such file or directory

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
1916it [00:05, 329.72it/s]
load_f1 = 0.6094890510948906
real_f1 = 0.6148282097649186
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, 202.05it/s]
```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

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

```

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

```

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

```
!mv *_dev.pt checkpoints/DBLP-ACM.pt
```

Step 3: run the matcher

```

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

```

Downloading: 100% 481/481 [00:00<00:00, 431kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 744kB/s]

Downloading: 100% 456k/456k [00:01<00:00, 400kB/s]

Downloading: 100% 501M/501M [00:10<00:00, 49.8MB/s]

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.564807653427124
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4297533631324768
step: 20, loss: 0.5516487956047058
step: 30, loss: 0.28701257705688477
step: 40, loss: 0.18429380655288696
step: 50, loss: 0.2449190467596054
step: 60, loss: 0.09337695688009262
step: 70, loss: 0.08734926581382751
step: 80, loss: 0.19211949408054352
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.2074487805366516
step: 100, loss: 0.17801129817962646
step: 110, loss: 0.06487247347831726
step: 120, loss: 0.0618782564997673
step: 130, loss: 0.08677094429731369
step: 140, loss: 0.015024441294372082
step: 150, loss: 0.12715472280979156
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 160, loss: 0.003943971823900938
step: 170, loss: 0.26892757415771484
step: 180, loss: 0.039002206176519394
step: 190, loss: 0.01385321095585823
step: 200, loss: 0.042406171560287476
step: 210, loss: 0.05627509951591492
step: 220, loss: 0.19037365913391113
step: 230, loss: 0.01412101462483406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9508571428571428, f1=0.9541284403669725,
best_f1=0.9541284403669725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025163467973470688
step: 10, loss: 0.165103942155838
step: 20, loss: 0.02260114625096321
step: 30, loss: 0.021092917770147324
step: 40, loss: 0.019963832572102547
step: 50, loss: 0.010240490548312664
step: 60, loss: 0.012220365926623344
step: 70, loss: 0.034861549735069275
step: 80, loss: 0.002331983530893922
step: 90, loss: 0.04429050534963608
step: 100, loss: 0.005002077668905258
step: 110, loss: 0.031325533986091614
step: 120, loss: 0.04094771668314934
step: 130, loss: 0.065371073782444
step: 140, loss: 0.010063597932457924
step: 150, loss: 0.10513950884342194
step: 160, loss: 0.016361655667424202
step: 170, loss: 0.012331683188676834
step: 180, loss: 0.02041245624423027
step: 190, loss: 0.08484108000993729
step: 200, loss: 0.0391845740377903
step: 210, loss: 0.021506156772375107
step: 220, loss: 0.01843593269586563
step: 230, loss: 0.0025693541392683983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9696969696969697, f1=0.9636363636363636,
best_f1=0.9636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023531407117843628
step: 10, loss: 0.00762079656124115

```

```

step: 20, loss: 0.01758972369134426
step: 30, loss: 0.003950200974941254
step: 40, loss: 0.03516264259815216
step: 50, loss: 0.08332644402980804
step: 60, loss: 0.02813202701508999
step: 70, loss: 0.018717166036367416
step: 80, loss: 0.10557029396295547
step: 90, loss: 0.03813569247722626
step: 100, loss: 0.005061613395810127
step: 110, loss: 0.008535685017704964
step: 120, loss: 0.0027831948827952147
step: 130, loss: 0.004577309358865023
step: 140, loss: 0.004638336598873138
step: 150, loss: 0.098726287484169
step: 160, loss: 0.017785558477044106
step: 170, loss: 0.0028247677255421877
step: 180, loss: 0.026518139988183975
step: 190, loss: 0.11079105734825134
step: 200, loss: 0.008576392196118832
step: 210, loss: 0.002120482036843896
step: 220, loss: 0.036752328276634216
step: 230, loss: 0.013797011226415634
/usr/local/lib/python3.7/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=0.9659863945578231, f1=0.9737742303306728,
best_f1=0.9636363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041070323437452316
step: 10, loss: 0.006484880577772856
step: 20, loss: 0.0055948542430996895
step: 30, loss: 0.010280396789312363
step: 40, loss: 0.022333214059472084
step: 50, loss: 0.008418474346399307
step: 60, loss: 0.011684635654091835
step: 70, loss: 0.02109358087182045
step: 80, loss: 0.022980058565735817
step: 90, loss: 0.005322948098182678
step: 100, loss: 0.0033201107289642096
step: 110, loss: 0.002198008121922612
step: 120, loss: 0.00890409667044878
step: 130, loss: 0.007903559133410454
step: 140, loss: 0.0010652905330061913
step: 150, loss: 0.0012248086277395487
step: 160, loss: 0.0019624752458184958
step: 170, loss: 0.0009577552555128932

```

```

step: 180, loss: 0.06365686655044556
step: 190, loss: 0.0027821743860840797
step: 200, loss: 0.006922677159309387
step: 210, loss: 0.07813326269388199
step: 220, loss: 0.0011374797904863954
step: 230, loss: 0.04019227996468544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9738933030646991, f1=0.971297359357061,
best_f1=0.971297359357061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002504336182028055
step: 10, loss: 0.003056056797504425
step: 20, loss: 0.004320717882364988
step: 30, loss: 0.0021697706542909145
step: 40, loss: 0.0036884527653455734
step: 50, loss: 0.0017627314664423466
step: 60, loss: 0.047404881566762924
step: 70, loss: 0.004833359736949205
step: 80, loss: 0.11772596836090088
step: 90, loss: 0.04739479348063469
step: 100, loss: 0.0007015002775005996
step: 110, loss: 0.0006441136356443167
step: 120, loss: 0.0014097357634454966
step: 130, loss: 0.0005535815143957734
step: 140, loss: 0.0014926670119166374
step: 150, loss: 0.01321756187826395
step: 160, loss: 0.018672820180654526
step: 170, loss: 0.002490613842383027
step: 180, loss: 0.0017940745456144214
step: 190, loss: 0.00775150815024972
step: 200, loss: 0.003136423183605075
step: 210, loss: 0.03737763315439224
step: 220, loss: 0.001589434570632875
step: 230, loss: 0.14169064164161682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9797752808988766, f1=0.9715585893060296,
best_f1=0.9715585893060296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035216426476836205
step: 10, loss: 0.0019380302401259542
step: 20, loss: 0.0019281900022178888
step: 30, loss: 0.0006112235714681447
step: 40, loss: 0.00046860743896104395
step: 50, loss: 0.0002831749734468758
step: 60, loss: 0.0012402479769662023
step: 70, loss: 0.13195671141147614
step: 80, loss: 0.002074070740491152
step: 90, loss: 0.004643748514354229
step: 100, loss: 0.0017221702728420496
step: 110, loss: 0.013963887467980385
step: 120, loss: 0.0007448501419275999
step: 130, loss: 0.06179412081837654
step: 140, loss: 0.0004965451662428677
step: 150, loss: 0.002072138013318181
step: 160, loss: 0.009962515905499458
step: 170, loss: 0.0010956812184304
step: 180, loss: 0.00040342361899092793
step: 190, loss: 0.0014037879882380366
step: 200, loss: 0.0064537543803453445
step: 210, loss: 0.006307034753262997
step: 220, loss: 0.0138021195307374
step: 230, loss: 0.017443224787712097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831271091113611, f1=0.9719416386083053,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015658042393624783
step: 10, loss: 0.0020954585634171963
step: 20, loss: 0.0008476717630401254
step: 30, loss: 0.0007848941022530198
step: 40, loss: 0.005675529129803181
step: 50, loss: 0.012529068626463413

```

```

step: 60, loss: 0.00042337950435467064
step: 70, loss: 0.0007942054653540254
step: 80, loss: 0.0007307685445994139
step: 90, loss: 0.0009928514482453465
step: 100, loss: 0.0012089596129953861
step: 110, loss: 0.0011727388482540846
step: 120, loss: 0.0008690867689438164
step: 130, loss: 0.0019767466001212597
step: 140, loss: 0.0002678185992408544
step: 150, loss: 0.005438077729195356
step: 160, loss: 0.0009559248574078083
step: 170, loss: 0.006117900367826223
step: 180, loss: 0.0023720362223684788
step: 190, loss: 0.005966739729046822
step: 200, loss: 0.000903647392988205
step: 210, loss: 0.0007868417305871844
step: 220, loss: 0.0013901495840400457
step: 230, loss: 0.004406806081533432
/usr/local/lib/python3.7/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=0.9785794813979707, f1=0.9761092150170648,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002818965818732977
step: 10, loss: 0.01506307814270258
step: 20, loss: 0.08916109800338745
step: 30, loss: 0.07433689385652542
step: 40, loss: 0.0027059258427470922
step: 50, loss: 0.0010526090627536178
step: 60, loss: 0.001860043965280056
step: 70, loss: 0.011505392380058765
step: 80, loss: 0.004398331977427006
step: 90, loss: 0.0005495777586475015
step: 100, loss: 0.0007606637082062662
step: 110, loss: 0.0012862824369221926
step: 120, loss: 0.0009630063432268798
step: 130, loss: 0.0011091248597949743
step: 140, loss: 0.006585684604942799
step: 150, loss: 0.2742387652397156
step: 160, loss: 0.0022014500573277473
step: 170, loss: 0.009584433399140835
step: 180, loss: 0.025723934173583984
step: 190, loss: 0.0009995547588914633
step: 200, loss: 0.0327884815633297
step: 210, loss: 0.004645452369004488

```

```

step: 220, loss: 0.0014207020867615938
step: 230, loss: 0.000854150508530438
/usr/local/lib/python3.7/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=0.9807037457434733, f1=0.9713631156930126,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019601471722126007
step: 10, loss: 0.0008025029092095792
step: 20, loss: 0.0009756376384757459
step: 30, loss: 0.005121797788888216
step: 40, loss: 0.001956250751391053
step: 50, loss: 0.0014685624046251178
step: 60, loss: 0.0007464203517884016
step: 70, loss: 0.03822178393602371
step: 80, loss: 0.00020848905842285603
step: 90, loss: 0.08581572026014328
step: 100, loss: 0.00038485968252643943
step: 110, loss: 0.0002947227912954986
step: 120, loss: 0.025172313675284386
step: 130, loss: 0.001154389581643045
step: 140, loss: 0.005972547456622124
step: 150, loss: 0.0010747804772108793
step: 160, loss: 0.070456862449646
step: 170, loss: 0.0005606524064205587
step: 180, loss: 0.0007580043748021126
step: 190, loss: 0.0635673776268959
step: 200, loss: 0.011444580741226673
step: 210, loss: 0.03308788314461708
step: 220, loss: 0.0009469327633269131
step: 230, loss: 0.0005362754454836249
/usr/local/lib/python3.7/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=0.9775784753363228, f1=0.9739524348810873,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010215967195108533
step: 10, loss: 0.0008151897927746177
step: 20, loss: 0.001282989513128996
step: 30, loss: 0.00029959692619740963
step: 40, loss: 0.0014232629910111427
step: 50, loss: 0.001286880113184452

```

```

step: 60, loss: 0.0006464797770604491
step: 70, loss: 0.004603639245033264
step: 80, loss: 0.0011260568862780929
step: 90, loss: 0.00018950965022668242
step: 100, loss: 0.0002290646079927683
step: 110, loss: 0.002061137929558754
step: 120, loss: 0.0002531330392230302
step: 130, loss: 0.00038846596726216376
step: 140, loss: 0.00022661649563815445
step: 150, loss: 0.000284312991425395
step: 160, loss: 8.649995288578793e-05
step: 170, loss: 0.0002630456583574414
step: 180, loss: 0.0005906244041398168
step: 190, loss: 0.0005099994014017284
step: 200, loss: 0.001117880456149578
step: 210, loss: 0.0009629758424125612
step: 220, loss: 0.04909569397568703
step: 230, loss: 0.00046504332567565143
/usr/local/lib/python3.7/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=0.9829351535836178, f1=0.9726651480637813,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006194582674652338
step: 10, loss: 0.001030131708830595
step: 20, loss: 0.0010360486339777708
step: 30, loss: 0.0006004552124068141
step: 40, loss: 9.383722499478608e-05
step: 50, loss: 0.0004508741549216211
step: 60, loss: 0.00922320131212473
step: 70, loss: 0.0016479501500725746
step: 80, loss: 0.0005920950789004564
step: 90, loss: 0.16390947997570038
step: 100, loss: 0.0005805248511023819
step: 110, loss: 0.0010821628384292126
step: 120, loss: 0.0003421300498303026
step: 130, loss: 0.00022977974731475115
step: 140, loss: 0.01005602814257145
step: 150, loss: 0.0004406201187521219
step: 160, loss: 0.0018194284057244658
step: 170, loss: 0.0013173818588256836
step: 180, loss: 0.001858262112364173
step: 190, loss: 0.0005619923467747867
step: 200, loss: 0.0013149534352123737
step: 210, loss: 0.0009859215933829546

```



```

step: 220, loss: 0.0006354115321300924
step: 230, loss: 0.0005688787787221372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841986455981941, f1=0.9727891156462585,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000608264293987304
step: 10, loss: 0.0001611778570804745
step: 20, loss: 0.011668316088616848
step: 30, loss: 0.0026943201664835215
step: 40, loss: 0.00042024313006550074
step: 50, loss: 0.00044814785360358655
step: 60, loss: 0.0005842758691869676
step: 70, loss: 0.0002570826036389917
step: 80, loss: 0.0002353664895053953
step: 90, loss: 0.0005431402241811156
step: 100, loss: 0.0002453806810081005
step: 110, loss: 0.00040391285438090563
step: 120, loss: 0.0007227911846712232
step: 130, loss: 0.0004383932682685554
step: 140, loss: 0.0011441153474152088
step: 150, loss: 0.0004767244099639356
step: 160, loss: 0.0018998971208930016
step: 170, loss: 0.00170269759837538
step: 180, loss: 0.0017236738931387663
step: 190, loss: 0.0020621786825358868
step: 200, loss: 0.0003198266203980893
step: 210, loss: 0.060666799545288086
step: 220, loss: 0.0017970781773328781
step: 230, loss: 0.004527405370026827
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9740112994350283,
best_f1=0.9727891156462585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008832024759612978
step: 10, loss: 0.0009736130014061928

```

```

step: 20, loss: 0.0023959174286574125
step: 30, loss: 9.26295542740263e-05
step: 40, loss: 0.0008596126572228968
step: 50, loss: 0.0016392734833061695
step: 60, loss: 0.000722908356692642
step: 70, loss: 0.005636018700897694
step: 80, loss: 0.0006410395726561546
step: 90, loss: 0.0005871693720109761
step: 100, loss: 0.0009259807993657887
step: 110, loss: 0.0025258525274693966
step: 120, loss: 0.0011920843971893191
step: 130, loss: 0.0015509374206885695
step: 140, loss: 0.0004262958827894181
step: 150, loss: 0.00024223688524216413
step: 160, loss: 0.0037554092705249786
step: 170, loss: 0.00031139777274802327
step: 180, loss: 0.005453118123114109
step: 190, loss: 0.0002419169177301228
step: 200, loss: 7.047780673019588e-05
step: 210, loss: 0.0008976193494163454
step: 220, loss: 0.0001864700752776116
step: 230, loss: 0.0004554332699626684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842696629213483, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011942796845687553
step: 10, loss: 0.00032877232297323644
step: 20, loss: 0.00017926222062669694
step: 30, loss: 0.000349063629982993
step: 40, loss: 0.00024264983949251473
step: 50, loss: 0.000185574492206797
step: 60, loss: 0.00020097869855817407
step: 70, loss: 0.0003933871630579233
step: 80, loss: 0.0003563673817552626
step: 90, loss: 0.0019739691633731127
step: 100, loss: 0.0004608622402884066
step: 110, loss: 0.0010405495995655656
step: 120, loss: 4.047348920721561e-05
step: 130, loss: 0.002681075595319271

```

```

step: 140, loss: 0.0004592329205479473
step: 150, loss: 0.0007139876252040267
step: 160, loss: 0.00020260378369130194
step: 170, loss: 0.0010316353291273117
step: 180, loss: 0.0005823852261528373
step: 190, loss: 0.00010948726412607357
step: 200, loss: 0.0003562843776308
step: 210, loss: 0.0002791859151329845
step: 220, loss: 0.00023553581559099257
step: 230, loss: 0.0016701024724170566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010422329418361187
step: 10, loss: 0.0001536373601993546
step: 20, loss: 0.001170448842458427
step: 30, loss: 0.00026711318059824407
step: 40, loss: 0.0001197649835376069
step: 50, loss: 7.959020877024159e-05
step: 60, loss: 0.010459567420184612
step: 70, loss: 0.0016753117088228464
step: 80, loss: 0.00026206718757748604
step: 90, loss: 0.00015918178542051464
step: 100, loss: 7.157595246098936e-05
step: 110, loss: 0.00015424737648572773
step: 120, loss: 0.033568523824214935
step: 130, loss: 0.0005937325768172741
step: 140, loss: 0.0018968931399285793
step: 150, loss: 0.00017192687664646655
step: 160, loss: 0.012257004156708717
step: 170, loss: 5.33180900674779e-05
step: 180, loss: 0.00020473050244618207
step: 190, loss: 0.0035951808094978333
step: 200, loss: 0.0017232151003554463
step: 210, loss: 0.009973814710974693
step: 220, loss: 0.0004325972986407578
step: 230, loss: 0.00011987981997663155
/usr/local/lib/python3.7/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=0.9832026875699889, f1=0.9764837625979844,
best_f1=0.9764837625979844
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.")
2473it [00:17, 142.44it/s]
load_f1 = 0.9854748603351955
real_f1 = 0.9821029082774049
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, 136.61it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
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: 0.7208261489868164
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41731059551239014
step: 20, loss: 0.3506248891353607
step: 30, loss: 0.3948271572589874
step: 40, loss: 0.19617338478565216
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.5094064474105835
step: 60, loss: 0.1623222678899765
step: 70, loss: 0.18014772236347198
step: 80, loss: 0.11208117753267288
step: 90, loss: 0.13906855881214142
step: 100, loss: 0.47078293561935425
step: 110, loss: 0.11808696389198303
step: 120, loss: 0.23722636699676514
step: 130, loss: 0.16298383474349976
step: 140, loss: 0.32329073548316956
step: 150, loss: 0.10458291321992874
step: 160, loss: 0.09984090924263
step: 170, loss: 0.04655781388282776
step: 180, loss: 0.08482000976800919
step: 190, loss: 0.041054390370845795
step: 200, loss: 0.1743331104516983
step: 210, loss: 0.1490347534418106
step: 220, loss: 0.04531518369913101
step: 230, loss: 0.08847472816705704
step: 240, loss: 0.03171071037650108
step: 250, loss: 0.06659876555204391
step: 260, loss: 0.16257883608341217
step: 270, loss: 0.37339621782302856
step: 280, loss: 0.06739874184131622

```

```

step: 290, loss: 0.10938166081905365
step: 300, loss: 0.05446028709411621
step: 310, loss: 0.1398921012878418
step: 320, loss: 0.10324323177337646
step: 330, loss: 0.07535836845636368
step: 340, loss: 0.3066200017929077
step: 350, loss: 0.07252293825149536
step: 360, loss: 0.03223886713385582
step: 370, loss: 0.07101593911647797
step: 380, loss: 0.1638464331626892
step: 390, loss: 0.015040051192045212
step: 400, loss: 0.08513890951871872
step: 410, loss: 0.2349909543991089
step: 420, loss: 0.08424166589975357
step: 430, loss: 0.0690818727016449
step: 440, loss: 0.049499280750751495
step: 450, loss: 0.040630124509334564
step: 460, loss: 0.016700247302651405
step: 470, loss: 0.060196541249752045
step: 480, loss: 0.100895456969738
step: 490, loss: 0.15076297521591187
step: 500, loss: 0.09051138907670975
step: 510, loss: 0.10176962614059448
step: 520, loss: 0.11244010925292969
step: 530, loss: 0.06710643321275711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9298245614035088, f1=0.9297597042513862,
best_f1=0.9297597042513862
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14394813776016235
step: 10, loss: 0.04738503694534302
step: 20, loss: 0.018451634794473648
step: 30, loss: 0.01640971750020981
step: 40, loss: 0.12205512821674347
step: 50, loss: 0.18248726427555084
step: 60, loss: 0.01786869764328003
step: 70, loss: 0.01244441233575344
step: 80, loss: 0.022209780290722847
step: 90, loss: 0.022022048011422157
step: 100, loss: 0.17162545025348663

```

```
step: 110, loss: 0.01043783314526081
step: 120, loss: 0.01848393864929676
step: 130, loss: 0.005996183957904577
step: 140, loss: 0.15478961169719696
step: 150, loss: 0.019294172525405884
step: 160, loss: 0.035602349787950516
step: 170, loss: 0.0525505468249321
step: 180, loss: 0.008947753347456455
step: 190, loss: 0.0031506691593676805
step: 200, loss: 0.12027031183242798
step: 210, loss: 0.034887656569480896
step: 220, loss: 0.006670203525573015
step: 230, loss: 0.06820300966501236
step: 240, loss: 0.002699154894798994
step: 250, loss: 0.054259661585092545
step: 260, loss: 0.027138935402035713
step: 270, loss: 0.050852108746767044
step: 280, loss: 0.06645364314317703
step: 290, loss: 0.04249089956283569
step: 300, loss: 0.027695786207914352
step: 310, loss: 0.05553487315773964
step: 320, loss: 0.04170946031808853
step: 330, loss: 0.17423301935195923
step: 340, loss: 0.21799908578395844
step: 350, loss: 0.0019406683277338743
step: 360, loss: 0.12220408022403717
step: 370, loss: 0.027679244056344032
step: 380, loss: 0.11994452029466629
step: 390, loss: 0.004657236859202385
step: 400, loss: 0.0544724203646183
step: 410, loss: 0.02133912965655327
step: 420, loss: 0.06719934940338135
step: 430, loss: 0.11130338162183762
step: 440, loss: 0.021950673311948776
step: 450, loss: 0.05013292655348778
step: 460, loss: 0.047421567142009735
step: 470, loss: 0.06730484217405319
step: 480, loss: 0.05901051312685013
step: 490, loss: 0.05787115544080734
step: 500, loss: 0.0036645501386374235
step: 510, loss: 0.049524009227752686
step: 520, loss: 0.30406710505485535
step: 530, loss: 0.06250070780515671
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9371584699453551, f1=0.9380127620783956,
best_f1=0.9380127620783956
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14033947885036469
step: 10, loss: 0.2750968933105469
step: 20, loss: 0.06849890202283859
step: 30, loss: 0.0605732686817646
step: 40, loss: 0.10058233141899109
step: 50, loss: 0.0027491494547575712
step: 60, loss: 0.00934015866369009
step: 70, loss: 0.014359411783516407
step: 80, loss: 0.03576921671628952
step: 90, loss: 0.03545499965548515
step: 100, loss: 0.05348643660545349
step: 110, loss: 0.03287656977772713
step: 120, loss: 0.14694203436374664
step: 130, loss: 0.12623344361782074
step: 140, loss: 0.02656704932451248
step: 150, loss: 0.020757514983415604
step: 160, loss: 0.0023890468291938305
step: 170, loss: 0.01026911474764347
step: 180, loss: 0.0034937048330903053
step: 190, loss: 0.017087379470467567
step: 200, loss: 0.09982658922672272
step: 210, loss: 0.023267868906259537
step: 220, loss: 0.12220954895019531
step: 230, loss: 0.04161111265420914
step: 240, loss: 0.03047804720699787
step: 250, loss: 0.02804487757384777
step: 260, loss: 0.02608213573694229
step: 270, loss: 0.004021290689706802
step: 280, loss: 0.0009764556307345629
step: 290, loss: 0.0034305446315556765
step: 300, loss: 0.10821763426065445
step: 310, loss: 0.07563499361276627
step: 320, loss: 0.0441729910671711
step: 330, loss: 0.003701006295159459
step: 340, loss: 0.0022928190883249044
step: 350, loss: 0.12326404452323914
step: 360, loss: 0.02312406152486801
step: 370, loss: 0.07803698629140854
step: 380, loss: 0.0799250602722168
step: 390, loss: 0.021550772711634636
step: 400, loss: 0.16554629802703857

```

```

step: 410, loss: 0.07270729541778564
step: 420, loss: 0.010040998458862305
step: 430, loss: 0.01887296326458454
step: 440, loss: 0.2091255486011505
step: 450, loss: 0.012951460666954517
step: 460, loss: 0.11784528940916061
step: 470, loss: 0.21105599403381348
step: 480, loss: 0.2985934615135193
step: 490, loss: 0.02723541110754013
step: 500, loss: 0.03266235813498497
step: 510, loss: 0.008421999402344227
step: 520, loss: 0.06277871131896973
step: 530, loss: 0.00641339598223567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9421028253821214, f1=0.9255663430420712,
best_f1=0.9255663430420712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007226517889648676
step: 10, loss: 0.004156900104135275
step: 20, loss: 0.007309845648705959
step: 30, loss: 0.040406350046396255
step: 40, loss: 0.005104545503854752
step: 50, loss: 0.07995985448360443
step: 60, loss: 0.004202755633741617
step: 70, loss: 0.030630668625235558
step: 80, loss: 0.1514783650636673
step: 90, loss: 0.04815511777997017
step: 100, loss: 0.0020245041232556105
step: 110, loss: 0.08905773609876633
step: 120, loss: 0.03515656664967537
step: 130, loss: 0.05000946670770645
step: 140, loss: 0.0894552618265152
step: 150, loss: 0.0030990445520728827
step: 160, loss: 0.04989860579371452
step: 170, loss: 0.01856495626270771
step: 180, loss: 0.005737274885177612
step: 190, loss: 0.06337582319974899
step: 200, loss: 0.07872481644153595
step: 210, loss: 0.0016393272671848536
step: 220, loss: 0.005218889564275742

```

```

step: 230, loss: 0.07662505656480789
step: 240, loss: 0.019782476127147675
step: 250, loss: 0.08243149518966675
step: 260, loss: 0.0014077534433454275
step: 270, loss: 0.07619298994541168
step: 280, loss: 0.005486281588673592
step: 290, loss: 0.005566947627812624
step: 300, loss: 0.003043434116989374
step: 310, loss: 0.0037165842950344086
step: 320, loss: 0.05392272397875786
step: 330, loss: 0.042679060250520706
step: 340, loss: 0.01390956062823534
step: 350, loss: 0.04958617314696312
step: 360, loss: 0.15008564293384552
step: 370, loss: 0.005249158013612032
step: 380, loss: 0.0036097464617341757
step: 390, loss: 0.00022295974486041814
step: 400, loss: 0.016498425975441933
step: 410, loss: 0.0018439082195982337
step: 420, loss: 0.024395478889346123
step: 430, loss: 0.00728612719103694
step: 440, loss: 0.0026121041737496853
step: 450, loss: 0.013016235083341599
step: 460, loss: 0.05622696131467819
step: 470, loss: 0.0014320939080789685
step: 480, loss: 0.05490565672516823
step: 490, loss: 0.01720457524061203
step: 500, loss: 0.09866756200790405
step: 510, loss: 0.2005043476819992
step: 520, loss: 0.029065165668725967
step: 530, loss: 0.13264881074428558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9474671669793621, f1=0.9400278940027894,
best_f1=0.9400278940027894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013754062820225954
step: 10, loss: 0.0033352875616401434
step: 20, loss: 0.009858457371592522
step: 30, loss: 0.06311589479446411
step: 40, loss: 0.003096937667578459

```

step: 50, loss: 0.002464870922267437
step: 60, loss: 0.04456410929560661
step: 70, loss: 0.0006936682621017098
step: 80, loss: 0.0037226404529064894
step: 90, loss: 0.0028076793532818556
step: 100, loss: 0.014474115334451199
step: 110, loss: 0.006425668951123953
step: 120, loss: 0.06517960876226425
step: 130, loss: 0.007255617529153824
step: 140, loss: 0.0029722116887569427
step: 150, loss: 0.07925143837928772
step: 160, loss: 0.01961151324212551
step: 170, loss: 0.0753280371427536
step: 180, loss: 0.008197293616831303
step: 190, loss: 0.114530049264431
step: 200, loss: 0.009241798892617226
step: 210, loss: 0.0009448539931327105
step: 220, loss: 0.0627857968211174
step: 230, loss: 0.0034928449895232916
step: 240, loss: 0.014313950203359127
step: 250, loss: 0.10606885701417923
step: 260, loss: 0.015109394676983356
step: 270, loss: 0.020629480481147766
step: 280, loss: 0.007969274185597897
step: 290, loss: 0.002716844668611884
step: 300, loss: 0.01900378428399563
step: 310, loss: 0.004397440701723099
step: 320, loss: 0.01616804488003254
step: 330, loss: 0.020200073719024658
step: 340, loss: 0.004001704044640064
step: 350, loss: 0.007328803185373545
step: 360, loss: 0.0008651597891002893
step: 370, loss: 0.00368562713265419
step: 380, loss: 0.0018244166858494282
step: 390, loss: 0.001634144107811153
step: 400, loss: 0.07369483262300491
step: 410, loss: 0.017978554591536522
step: 420, loss: 0.10687561333179474
step: 430, loss: 0.11533123254776001
step: 440, loss: 0.004594144411385059
step: 450, loss: 0.15777041018009186
step: 460, loss: 0.01463061198592186
step: 470, loss: 0.02411031723022461
step: 480, loss: 0.022779151797294617
step: 490, loss: 0.0374239943921566
step: 500, loss: 0.016060572117567062
step: 510, loss: 0.0465354360640049
step: 520, loss: 0.0535159595310688

```
step: 530, loss: 0.03471904620528221
/usr/local/lib/python3.7/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=0.944704779756326, f1=0.9320754716981132,
best_f1=0.9400278940027894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0393865592777729
step: 10, loss: 0.013587665744125843
step: 20, loss: 0.014565971679985523
step: 30, loss: 0.001538545242510736
step: 40, loss: 0.01110862661153078
step: 50, loss: 0.0009594039293006063
step: 60, loss: 0.002552961464971304
step: 70, loss: 0.004559722729027271
step: 80, loss: 6.457351992139593e-05
step: 90, loss: 0.004331469070166349
step: 100, loss: 0.02410808578133583
step: 110, loss: 0.030129902064800262
step: 120, loss: 0.04276247322559357
step: 130, loss: 0.024874577298760414
step: 140, loss: 0.0013140342198312283
step: 150, loss: 0.00030014297226443887
step: 160, loss: 0.003322798293083906
step: 170, loss: 0.0018159158062189817
step: 180, loss: 0.005836889613419771
step: 190, loss: 0.13357438147068024
step: 200, loss: 0.007295504678040743
step: 210, loss: 0.00150210689753294
step: 220, loss: 0.07401008158922195
step: 230, loss: 0.006890340242534876
step: 240, loss: 0.016233572736382484
step: 250, loss: 0.0718943327665329
step: 260, loss: 0.013742375187575817
step: 270, loss: 0.008799106813967228
step: 280, loss: 0.08524863421916962
step: 290, loss: 0.01959274709224701
step: 300, loss: 0.007681285496801138
step: 310, loss: 0.10221588611602783
step: 320, loss: 0.0012568809324875474
step: 330, loss: 0.0031296799425035715
step: 340, loss: 0.0029651192016899586
step: 350, loss: 0.026075217872858047
step: 360, loss: 0.025676734745502472
step: 370, loss: 0.0010926842223852873
step: 380, loss: 0.00011530805932125077
```

```

step: 390, loss: 0.002187141450121999
step: 400, loss: 0.03437718376517296
step: 410, loss: 6.916873098816723e-05
step: 420, loss: 0.12723681330680847
step: 430, loss: 0.0066863554529845715
step: 440, loss: 0.0028910706751048565
step: 450, loss: 0.050317294895648956
step: 460, loss: 0.0015198104083538055
step: 470, loss: 0.001197111327201128
step: 480, loss: 0.0008292206330224872
step: 490, loss: 0.009356536902487278
step: 500, loss: 0.013378445990383625
step: 510, loss: 0.05670247599482536
step: 520, loss: 0.007293465081602335
step: 530, loss: 0.0222573671489954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9500924214417744, f1=0.9390018484288354,
best_f1=0.9390018484288354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005579730030149221
step: 10, loss: 0.0005409774603322148
step: 20, loss: 0.0018320393282920122
step: 30, loss: 0.01558644138276577
step: 40, loss: 0.0003778492391575128
step: 50, loss: 0.022419700399041176
step: 60, loss: 0.0018114240374416113
step: 70, loss: 0.002037315396592021
step: 80, loss: 0.0030162313487380743
step: 90, loss: 0.0007006244268268347
step: 100, loss: 0.013811304233968258
step: 110, loss: 0.00016059317567851394
step: 120, loss: 0.00023018325737211853
step: 130, loss: 0.00032236508559435606
step: 140, loss: 0.0017035611672326922
step: 150, loss: 0.0012629976263269782
step: 160, loss: 0.00018433154036756605
step: 170, loss: 0.05402996763586998
step: 180, loss: 0.0467509999871254
step: 190, loss: 0.0038089591544121504
step: 200, loss: 0.00037447645445354283

```

```

step: 210, loss: 0.0012316345237195492
step: 220, loss: 3.7698180676670745e-05
step: 230, loss: 0.00020946857694070786
step: 240, loss: 0.00016330975631717592
step: 250, loss: 0.00454458175227046
step: 260, loss: 0.010027243755757809
step: 270, loss: 0.0021431308705359697
step: 280, loss: 0.03013310208916664
step: 290, loss: 0.007711514830589294
step: 300, loss: 7.256653043441474e-05
step: 310, loss: 0.00017249178199563175
step: 320, loss: 0.018479540944099426
step: 330, loss: 0.003022597637027502
step: 340, loss: 0.00045875334762968123
step: 350, loss: 0.0019877790473401546
step: 360, loss: 0.034809648990631104
step: 370, loss: 0.015375983901321888
step: 380, loss: 0.008303419686853886
step: 390, loss: 0.0283404178917408
step: 400, loss: 0.01181887835264206
step: 410, loss: 0.13296343386173248
step: 420, loss: 0.005948410835117102
step: 430, loss: 0.007763804867863655
step: 440, loss: 0.049911487847566605
step: 450, loss: 0.007467741146683693
step: 460, loss: 0.0016000756295397878
step: 470, loss: 0.19112299382686615
step: 480, loss: 0.0021029645577073097
step: 490, loss: 0.026155643165111542
step: 500, loss: 0.0017544450238347054
step: 510, loss: 0.000314609264023602
step: 520, loss: 0.0010884609073400497
step: 530, loss: 0.0042085954919457436
/usr/local/lib/python3.7/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=0.9417840375586854, f1=0.9296037296037295,
best_f1=0.9390018484288354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004388396628201008
step: 10, loss: 0.003141577821224928
step: 20, loss: 0.0037167815025895834
step: 30, loss: 0.0026437253691256046
step: 40, loss: 0.002092626178637147
step: 50, loss: 0.0012820007978007197
step: 60, loss: 0.0008753429283387959

```

step: 70, loss: 0.04465221241116524
step: 80, loss: 0.000862916640471667
step: 90, loss: 0.01816745474934578
step: 100, loss: 0.012518583796918392
step: 110, loss: 0.0006112545961514115
step: 120, loss: 0.004962841048836708
step: 130, loss: 0.0019459922332316637
step: 140, loss: 0.04140333831310272
step: 150, loss: 0.001040967763401568
step: 160, loss: 0.007733880542218685
step: 170, loss: 0.28221267461776733
step: 180, loss: 0.006067198235541582
step: 190, loss: 0.004270344972610474
step: 200, loss: 0.039042290300130844
step: 210, loss: 0.12464630603790283
step: 220, loss: 0.027059121057391167
step: 230, loss: 0.04475095495581627
step: 240, loss: 0.009691379964351654
step: 250, loss: 0.011198081076145172
step: 260, loss: 0.0022747262846678495
step: 270, loss: 0.009748916141688824
step: 280, loss: 0.002491454128175974
step: 290, loss: 0.04669272527098656
step: 300, loss: 0.0014232132816687226
step: 310, loss: 0.0012874230742454529
step: 320, loss: 0.02429487742483616
step: 330, loss: 0.00024648697581142187
step: 340, loss: 0.006578438449651003
step: 350, loss: 0.00348775926977396
step: 360, loss: 0.009664276614785194
step: 370, loss: 0.07319650053977966
step: 380, loss: 0.0006762562552466989
step: 390, loss: 0.0480726957321167
step: 400, loss: 0.03292554244399071
step: 410, loss: 0.08295812457799911
step: 420, loss: 0.00014206110790837556
step: 430, loss: 0.02668185718357563
step: 440, loss: 0.003138133557513356
step: 450, loss: 0.0032398905605077744
step: 460, loss: 0.047794077545404434
step: 470, loss: 0.22053347527980804
step: 480, loss: 0.00161934329662472
step: 490, loss: 0.006064313463866711
step: 500, loss: 0.005178837571293116
step: 510, loss: 0.006676522549241781
step: 520, loss: 0.005488567985594273
step: 530, loss: 0.00031088723335415125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:


```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9505316689782709, f1=0.9362092703074805,
best_f1=0.9362092703074805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007449167314916849
step: 10, loss: 0.003372818697243929
step: 20, loss: 0.001305879675783217
step: 30, loss: 0.02887612208724022
step: 40, loss: 0.040868427604436874
step: 50, loss: 0.0011604288592934608
step: 60, loss: 0.009473457932472229
step: 70, loss: 0.0008031547185964882
step: 80, loss: 0.12628445029258728
step: 90, loss: 0.09673059731721878
step: 100, loss: 0.001978003652766347
step: 110, loss: 0.010408051311969757
step: 120, loss: 0.001772148534655571
step: 130, loss: 0.0005342814256437123
step: 140, loss: 0.002296506194397807
step: 150, loss: 0.006226508412510157
step: 160, loss: 0.0037704790011048317
step: 170, loss: 0.00043913419358432293
step: 180, loss: 0.022861963137984276
step: 190, loss: 0.0028275379445403814
step: 200, loss: 8.447823347523808e-05
step: 210, loss: 0.0005230656242929399
step: 220, loss: 0.000497045402880758
step: 230, loss: 0.004423648584634066
step: 240, loss: 0.0055889650247991085
step: 250, loss: 0.001241447520442307
step: 260, loss: 0.0024954488035291433
step: 270, loss: 0.02276911586523056
step: 280, loss: 0.010722186416387558
step: 290, loss: 0.0009675096953287721
step: 300, loss: 0.0004382047918625176
step: 310, loss: 0.14642341434955597
step: 320, loss: 0.15087954699993134
step: 330, loss: 0.0006147018284536898
step: 340, loss: 0.0934896394610405
step: 350, loss: 0.17930035293102264
step: 360, loss: 0.0012294645421206951

```

```
step: 370, loss: 0.003584421705454588
step: 380, loss: 0.011428819969296455
step: 390, loss: 0.001555754104629159
step: 400, loss: 0.004130814224481583
step: 410, loss: 0.0005765542155131698
step: 420, loss: 0.0011611268855631351
step: 430, loss: 0.03741292282938957
step: 440, loss: 0.005262494552880526
step: 450, loss: 0.028235413134098053
step: 460, loss: 0.00036313384771347046
step: 470, loss: 0.0006265139672905207
step: 480, loss: 0.0015486221527680755
step: 490, loss: 0.018901150673627853
step: 500, loss: 0.00677532609552145
step: 510, loss: 0.05661391094326973
step: 520, loss: 0.004965729545801878
step: 530, loss: 0.006924069952219725
/usr/local/lib/python3.7/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=0.9479981592268752, f1=0.9371847776249428,
best_f1=0.9362092703074805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004704564809799194
step: 10, loss: 0.0004879040934611112
step: 20, loss: 0.0013632085174322128
step: 30, loss: 0.0007825755164958537
step: 40, loss: 0.0013658016687259078
step: 50, loss: 0.0012252095621079206
step: 60, loss: 0.025947099551558495
step: 70, loss: 0.0006137277232483029
step: 80, loss: 0.002292510587722063
step: 90, loss: 0.0002593040117062628
step: 100, loss: 0.003682107664644718
step: 110, loss: 0.023849599063396454
step: 120, loss: 0.0008683076594024897
step: 130, loss: 0.00017448890139348805
step: 140, loss: 0.00014863621618133038
step: 150, loss: 0.006293043959885836
step: 160, loss: 0.0054657235741615295
step: 170, loss: 0.011356120929121971
step: 180, loss: 0.0071091726422309875
step: 190, loss: 0.002203435404226184
step: 200, loss: 0.0016988886054605246
step: 210, loss: 0.021792501211166382
step: 220, loss: 0.0018952953396365047
```

```

step: 230, loss: 0.0008552017970941961
step: 240, loss: 0.0008940143161453307
step: 250, loss: 0.0004619368992280215
step: 260, loss: 0.08864610642194748
step: 270, loss: 0.00027946149930357933
step: 280, loss: 0.008158288896083832
step: 290, loss: 0.0032208114862442017
step: 300, loss: 0.002687284490093589
step: 310, loss: 0.004007870331406593
step: 320, loss: 0.0016701847780495882
step: 330, loss: 0.012498252093791962
step: 340, loss: 0.0011204518377780914
step: 350, loss: 0.19554857909679413
step: 360, loss: 0.0005315595772117376
step: 370, loss: 0.014721297658979893
step: 380, loss: 0.005018214695155621
step: 390, loss: 0.0002803526585921645
step: 400, loss: 0.0016903612995520234
step: 410, loss: 0.0018310000887140632
step: 420, loss: 0.0012355716899037361
step: 430, loss: 0.0010239268885925412
step: 440, loss: 0.0001121004024753347
step: 450, loss: 0.0002918616228271276
step: 460, loss: 0.0010045368690043688
step: 470, loss: 0.0013465478550642729
step: 480, loss: 0.0005206966889090836
step: 490, loss: 0.0010589949088171124
step: 500, loss: 0.00704084150493145
step: 510, loss: 0.000405426457291469
step: 520, loss: 0.0731925442814827
step: 530, loss: 0.0012123910710215569
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 10: dev_f1=0.9486348912540491, f1=0.9363086936308694,
best_f1=0.9362092703074805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003322848060633987
step: 10, loss: 6.186791870277375e-05
step: 20, loss: 0.002459537470713258
step: 30, loss: 0.003253024537116289
step: 40, loss: 0.00024765764828771353
step: 50, loss: 0.0006245410768315196
step: 60, loss: 0.0020902405958622694
step: 70, loss: 0.0005046877195127308

```

```
step: 80, loss: 0.001182050909847021
step: 90, loss: 0.0046755229122936726
step: 100, loss: 0.0036526776384562254
step: 110, loss: 0.0011826095869764686
step: 120, loss: 0.0005640657036565244
step: 130, loss: 0.0010164575651288033
step: 140, loss: 0.00013457766908686608
step: 150, loss: 0.0022431646939367056
step: 160, loss: 0.001478402642533183
step: 170, loss: 0.025632472708821297
step: 180, loss: 4.249401536071673e-05
step: 190, loss: 0.0007028415566310287
step: 200, loss: 0.0016323686577379704
step: 210, loss: 0.00041493753087706864
step: 220, loss: 0.000784456089604646
step: 230, loss: 0.0003903216274920851
step: 240, loss: 0.0023136772215366364
step: 250, loss: 1.0736123840615619e-05
step: 260, loss: 0.0015577003359794617
step: 270, loss: 0.003952522296458483
step: 280, loss: 0.029112134128808975
step: 290, loss: 0.002114493167027831
step: 300, loss: 0.00014559787814505398
step: 310, loss: 0.0026044233236461878
step: 320, loss: 0.0004557106294669211
step: 330, loss: 1.0393474440206774e-05
step: 340, loss: 0.0029786594677716494
step: 350, loss: 0.0027704332023859024
step: 360, loss: 0.06981941312551498
step: 370, loss: 0.00033153328695334494
step: 380, loss: 0.0014335284940898418
step: 390, loss: 0.003069564001634717
step: 400, loss: 2.5773586457944475e-05
step: 410, loss: 0.0006621085340157151
step: 420, loss: 5.079658512840979e-05
step: 430, loss: 0.007336802314966917
step: 440, loss: 7.120834197849035e-05
step: 450, loss: 0.0007781851454637945
step: 460, loss: 0.01638297736644745
step: 470, loss: 0.0007420505280606449
step: 480, loss: 0.00021772057516500354
step: 490, loss: 0.00011347110557835549
step: 500, loss: 0.0007420906331390142
step: 510, loss: 0.0012199210468679667
step: 520, loss: 0.00013503464288078249
step: 530, loss: 0.12568077445030212
/usr/local/lib/python3.7/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=0.942095588235294, f1=0.9308924485125859,
best_f1=0.9362092703074805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004281776200514287
step: 10, loss: 0.0012871181825175881
step: 20, loss: 0.002552968682721257
step: 30, loss: 0.00013289068010635674
step: 40, loss: 0.0016085106180980802
step: 50, loss: 0.02164877951145172
step: 60, loss: 0.0004868795513175428
step: 70, loss: 0.0024702518712729216
step: 80, loss: 0.00020770962873939425
step: 90, loss: 0.0058720228262245655
step: 100, loss: 0.005314957350492477
step: 110, loss: 0.00021866150200366974
step: 120, loss: 0.00015647034160792828
step: 130, loss: 0.002647171728312969
step: 140, loss: 0.09045674651861191
step: 150, loss: 3.391972495592199e-05
step: 160, loss: 0.0006803278811275959
step: 170, loss: 4.878800245933235e-05
step: 180, loss: 0.0035886107943952084
step: 190, loss: 0.0005172966048121452
step: 200, loss: 0.000495883054099977
step: 210, loss: 0.00046491759712807834
step: 220, loss: 0.0004337886057328433
step: 230, loss: 0.006979544181376696
step: 240, loss: 0.0008853835752233863
step: 250, loss: 0.11993085592985153
step: 260, loss: 0.0008535758242942393
step: 270, loss: 0.0013048246037214994
step: 280, loss: 0.001876886817626655
step: 290, loss: 0.0036172522231936455
step: 300, loss: 0.0017959829419851303
step: 310, loss: 0.0069994074292480946
step: 320, loss: 0.03297688439488411
step: 330, loss: 0.0031590168364346027
step: 340, loss: 0.02382401004433632
step: 350, loss: 0.00011157608969369903
step: 360, loss: 0.001401923829689622
step: 370, loss: 0.0009576893644407392
step: 380, loss: 0.00027028092881664634
step: 390, loss: 0.009178532287478447
step: 400, loss: 0.0001351269893348217
step: 410, loss: 0.0004968939465470612

```

```

step: 420, loss: 0.003189110429957509
step: 430, loss: 0.0020538067910820246
step: 440, loss: 0.003033652435988188
step: 450, loss: 0.0010519670322537422
step: 460, loss: 0.0008796390029601753
step: 470, loss: 0.0034704485442489386
step: 480, loss: 0.013085935264825821
step: 490, loss: 0.0008689548121765256
step: 500, loss: 0.00033526282641105354
step: 510, loss: 0.003132549347355962
step: 520, loss: 0.000151230618939735
step: 530, loss: 0.00018578083836473525
/usr/local/lib/python3.7/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=0.9486461251167133, f1=0.9377323420074348,
best_f1=0.9362092703074805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001133263693191111
step: 10, loss: 0.0006768024177290499
step: 20, loss: 0.0014872490428388119
step: 30, loss: 1.152961249317741e-05
step: 40, loss: 0.037663884460926056
step: 50, loss: 0.00127887143753469
step: 60, loss: 0.00041768778464756906
step: 70, loss: 0.054923467338085175
step: 80, loss: 0.003084995551034808
step: 90, loss: 0.00036112265661358833
step: 100, loss: 0.02092961221933365
step: 110, loss: 0.0009852222865447402
step: 120, loss: 0.00014164713502395898
step: 130, loss: 0.0012103498447686434
step: 140, loss: 0.0008583373273722827
step: 150, loss: 0.01828853040933609
step: 160, loss: 0.002564270282164216
step: 170, loss: 0.1384160965681076
step: 180, loss: 0.0023322284687310457
step: 190, loss: 0.00010678703984012827
step: 200, loss: 0.0051544844172894955
step: 210, loss: 3.652410305221565e-05
step: 220, loss: 0.0008095118100754917
step: 230, loss: 0.002048541558906436
step: 240, loss: 0.0024416563101112843
step: 250, loss: 0.0012787358136847615
step: 260, loss: 0.001039651921018958
step: 270, loss: 0.11217743903398514

```

```

step: 280, loss: 0.000779110356234014
step: 290, loss: 0.0007774113910272717
step: 300, loss: 0.0012248557759448886
step: 310, loss: 0.0008499794639647007
step: 320, loss: 0.00274123577401042
step: 330, loss: 8.228297519963235e-05
step: 340, loss: 0.0022049483377486467
step: 350, loss: 0.0004926439141854644
step: 360, loss: 0.09814946353435516
step: 370, loss: 0.007040213327854872
step: 380, loss: 0.0005395157495513558
step: 390, loss: 0.006503280717879534
step: 400, loss: 0.0013353020185604692
step: 410, loss: 0.00017438259965274483
step: 420, loss: 0.0011005975538864732
step: 430, loss: 0.0007130011217668653
step: 440, loss: 0.00016306013276334852
step: 450, loss: 0.000571548705920577
step: 460, loss: 0.0009994667489081621
step: 470, loss: 0.004137189593166113
step: 480, loss: 0.0003222355153411627
step: 490, loss: 0.00011902500409632921
step: 500, loss: 0.0001747835922287777
step: 510, loss: 0.000722161668818444
step: 520, loss: 0.0001924919633893296
step: 530, loss: 0.0002198194561060518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9528698086794214, f1=0.9402501157943491,
best_f1=0.9402501157943491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007204197463579476
step: 10, loss: 0.0016325905453413725
step: 20, loss: 0.00047163182171061635
step: 30, loss: 0.006388308014720678
step: 40, loss: 0.00018245608953293413
step: 50, loss: 0.00518425926566124
step: 60, loss: 0.0008734029834158719
step: 70, loss: 0.0005567888729274273
step: 80, loss: 0.0004138282092753798
step: 90, loss: 0.00012637345935218036

```

```
step: 100, loss: 0.00037040296592749655
step: 110, loss: 0.0009235473116859794
step: 120, loss: 7.362765609286726e-05
step: 130, loss: 7.627743616467342e-05
step: 140, loss: 0.0297746192663908
step: 150, loss: 0.00017317431047558784
step: 160, loss: 0.00014273193664848804
step: 170, loss: 0.0003037089772988111
step: 180, loss: 0.000250600918661803
step: 190, loss: 0.002222080249339342
step: 200, loss: 6.129437679192051e-05
step: 210, loss: 0.00044026263640262187
step: 220, loss: 5.467875234899111e-05
step: 230, loss: 0.00037191997398622334
step: 240, loss: 0.0010174415074288845
step: 250, loss: 0.05283748358488083
step: 260, loss: 0.16557785868644714
step: 270, loss: 0.0002899138198699802
step: 280, loss: 0.0003960324975196272
step: 290, loss: 0.000303042063023895
step: 300, loss: 0.0003493853728286922
step: 310, loss: 0.020659197121858597
step: 320, loss: 0.00018148224626202136
step: 330, loss: 0.0012499872827902436
step: 340, loss: 0.00039357724017463624
step: 350, loss: 6.707790453219786e-05
step: 360, loss: 7.31443942640908e-05
step: 370, loss: 0.009429420344531536
step: 380, loss: 0.0022511433344334364
step: 390, loss: 0.0011115969391539693
step: 400, loss: 0.004819415509700775
step: 410, loss: 0.00020915466302540153
step: 420, loss: 0.04605463147163391
step: 430, loss: 0.000832880032248795
step: 440, loss: 0.006620470900088549
step: 450, loss: 0.0013910734560340643
step: 460, loss: 0.005519534461200237
step: 470, loss: 7.854776049498469e-05
step: 480, loss: 0.0015158277237787843
step: 490, loss: 0.0029906879644840956
step: 500, loss: 0.00870178360491991
step: 510, loss: 0.0011000687954947352
step: 520, loss: 0.00027010260964743793
step: 530, loss: 0.001754925586283207
/usr/local/lib/python3.7/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=0.9523368810735771, f1=0.9405803777061263,
```



```
best_f1=0.9402501157943491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019053411961067468
step: 10, loss: 0.0013015023432672024
step: 20, loss: 0.002361125312745571
step: 30, loss: 0.004930063616484404
step: 40, loss: 0.00021033968369010836
step: 50, loss: 0.0015587455127388239
step: 60, loss: 0.0005365770775824785
step: 70, loss: 0.0001267286716029048
step: 80, loss: 0.00026221247389912605
step: 90, loss: 0.0010297077242285013
step: 100, loss: 0.00012599927140399814
step: 110, loss: 0.0002062011626549065
step: 120, loss: 4.820443791686557e-05
step: 130, loss: 0.0002242157352156937
step: 140, loss: 0.00012093201075913385
step: 150, loss: 0.0001078929562936537
step: 160, loss: 0.00017071448382921517
step: 170, loss: 7.801932224538177e-05
step: 180, loss: 0.00015679745411034673
step: 190, loss: 0.0021559300366789103
step: 200, loss: 0.0021023331210017204
step: 210, loss: 0.00013819652667734772
step: 220, loss: 2.7226609745412134e-05
step: 230, loss: 0.01881333813071251
step: 240, loss: 9.927619248628616e-05
step: 250, loss: 0.0030895124655216932
step: 260, loss: 4.08029263780918e-05
step: 270, loss: 0.0011033324990421534
step: 280, loss: 3.366647069924511e-05
step: 290, loss: 0.0007265859167091548
step: 300, loss: 0.00016691013297531754
step: 310, loss: 0.04905768483877182
step: 320, loss: 0.00011304783402010798
step: 330, loss: 0.0002214450214523822
step: 340, loss: 0.00022872912813909352
step: 350, loss: 0.0044379751197993755
step: 360, loss: 0.00020460969244595617
step: 370, loss: 0.015705741941928864
step: 380, loss: 0.0004068720154464245
step: 390, loss: 4.06275357818231e-05
step: 400, loss: 0.004381339997053146
step: 410, loss: 0.0003863669408019632
step: 420, loss: 0.0013437814777716994
step: 430, loss: 0.0001170868199551478
```

```

step: 440, loss: 0.001018883311189711
step: 450, loss: 0.001118990476243198
step: 460, loss: 0.0008643961627967656
step: 470, loss: 6.778885290259495e-05
step: 480, loss: 0.0014984471490606666
step: 490, loss: 0.0014247136423364282
step: 500, loss: 0.007464051712304354
step: 510, loss: 0.00028430449310690165
step: 520, loss: 1.6763458916102536e-05
step: 530, loss: 0.00024087345809675753
/usr/local/lib/python3.7/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=0.9506976744186046, f1=0.9397031539888682,
best_f1=0.9402501157943491
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 [00:34, 167.61it/s]
load_f1 = 0.9534450651769087
real_f1 = 0.9496738117427772

```

```
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 142.82it/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: 0.43292754888534546
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=0.3218390804597701, 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: 0.39040783047676086
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
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: 0.410506933927536
/usr/local/lib/python3.7/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=0.2745098039215686, f1=0.3023255813953489,
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: 0.31688961386680603
/usr/local/lib/python3.7/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=0.30434782608695654, f1=0.2857142857142857,
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: 0.3122955858707428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.34210526315789475, f1=0.3, best_f1=0.3
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24305710196495056
/usr/local/lib/python3.7/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=0.2745098039215686, f1=0.35135135135135137, best_f1=0.3
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5130338668823242
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.4, f1=0.48888888888888893, best_f1=0.48888888888888893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31974244117736816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.46428571428571436, f1=0.5000000000000001,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25407281517982483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4905660377358491, f1=0.45161290322580644,
best_f1=0.45161290322580644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23102864623069763
/usr/local/lib/python3.7/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=0.44444444444444445, f1=0.5531914893617021,
best_f1=0.45161290322580644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24230292439460754
/usr/local/lib/python3.7/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=0.48148148148148157, f1=0.5945945945945946,

```

```

best_f1=0.45161290322580644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16386152803897858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5581395348837208, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09454485028982162
/usr/local/lib/python3.7/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=0.49122807017543857, f1=0.5945945945945946,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10643554478883743
/usr/local/lib/python3.7/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=0.49122807017543857, f1=0.5945945945945946,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16743804514408112
/usr/local/lib/python3.7/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=0.49122807017543857, f1=0.5945945945945946,
best_f1=0.5925925925925927
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, 102300.10it/s]
/usr/local/lib/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 = 0.5306122448979592
real_f1 = 0.49122807017543857
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.03it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

step: 0, loss: 0.586753249168396

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

step: 10, loss: 0.44319602847099304

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

step: 20, loss: 0.47736337780952454

step: 30, loss: 0.28039032220840454

```

step: 40, loss: 0.308490514755249
step: 50, loss: 0.6247422695159912
step: 60, loss: 0.4759730100631714
step: 70, loss: 0.4285849928855896
step: 80, loss: 0.6099539995193481
step: 90, loss: 0.4629555344581604
step: 100, loss: 0.4950568675994873
step: 110, loss: 0.5852223038673401
step: 120, loss: 0.43616223335266113
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.1552935540676117
step: 140, loss: 0.12030753493309021
step: 150, loss: 0.15223829448223114
step: 160, loss: 0.047423433512449265
step: 170, loss: 0.22106605768203735
step: 180, loss: 0.18525108695030212
step: 190, loss: 0.2663538157939911
step: 200, loss: 0.06445156037807465
step: 210, loss: 0.0856584683060646
step: 220, loss: 0.15595529973506927
step: 230, loss: 0.016394294798374176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9599999999999999, f1=0.9516483516483516,
best_f1=0.9516483516483516
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11499536782503128
step: 10, loss: 0.15516774356365204
step: 20, loss: 0.08532114326953888
step: 30, loss: 0.07934780418872833
step: 40, loss: 0.07924312353134155
step: 50, loss: 0.021161789074540138
step: 60, loss: 0.03149382397532463
step: 70, loss: 0.05812068283557892
step: 80, loss: 0.01916339062154293
step: 90, loss: 0.005978086963295937
step: 100, loss: 0.10731802880764008
step: 110, loss: 0.02839544042944908
step: 120, loss: 0.03897407278418541
step: 130, loss: 0.007195227313786745
step: 140, loss: 0.0024208843242377043

```

```

step: 150, loss: 0.06974537670612335
step: 160, loss: 0.014492515474557877
step: 170, loss: 0.007408624514937401
step: 180, loss: 0.011229554191231728
step: 190, loss: 0.040388599038124084
step: 200, loss: 0.047627631574869156
step: 210, loss: 0.017091281712055206
step: 220, loss: 0.027370022609829903
step: 230, loss: 0.039587512612342834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819819819819819, f1=0.9832026875699889,
best_f1=0.9832026875699889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006346203852444887
step: 10, loss: 0.004259107634425163
step: 20, loss: 0.0006831747014075518
step: 30, loss: 0.0028292040806263685
step: 40, loss: 0.0349222794175148
step: 50, loss: 0.05848206579685211
step: 60, loss: 0.07375676929950714
step: 70, loss: 0.016064418479800224
step: 80, loss: 0.07032843679189682
step: 90, loss: 0.049949366599321365
step: 100, loss: 0.0018331978935748339
step: 110, loss: 0.0034391360823065042
step: 120, loss: 0.003972066566348076
step: 130, loss: 0.003287253202870488
step: 140, loss: 0.10911202430725098
step: 150, loss: 0.045642200857400894
step: 160, loss: 0.008168739266693592
step: 170, loss: 0.01612565852701664
step: 180, loss: 0.01691019907593727
step: 190, loss: 0.005774686112999916
step: 200, loss: 0.03520599752664566
step: 210, loss: 0.004900924861431122
step: 220, loss: 0.002684788778424263
step: 230, loss: 0.019797202199697495
/usr/local/lib/python3.7/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=0.9819413092550789, f1=0.9820224719101124,
best_f1=0.9832026875699889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006988632492721081
step: 10, loss: 0.002059153513982892
step: 20, loss: 0.0027130786329507828
step: 30, loss: 0.0012119847815483809
step: 40, loss: 0.041948746889829636
step: 50, loss: 0.007926111109554768
step: 60, loss: 0.0030039455741643906
step: 70, loss: 0.0005891930195502937
step: 80, loss: 0.0005448558367788792
step: 90, loss: 0.0046300096437335014
step: 100, loss: 0.0012843949953094125
step: 110, loss: 0.0018033263040706515
step: 120, loss: 0.0016789407236501575
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 130, loss: 0.0067617096938192844
step: 140, loss: 0.0013453603023663163
step: 150, loss: 0.0044273198582232
step: 160, loss: 0.0022786634508520365
step: 170, loss: 0.023197844624519348
step: 180, loss: 0.04959931969642639
step: 190, loss: 0.001867721788585186
step: 200, loss: 0.009077069349586964
step: 210, loss: 0.0006450144574046135
step: 220, loss: 0.0007381623727269471
step: 230, loss: 0.0036790312733501196
/usr/local/lib/python3.7/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=0.9729119638826186, f1=0.9819819819819819,
best_f1=0.9832026875699889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014556723181158304
step: 10, loss: 0.005559121258556843
step: 20, loss: 0.002768662292510271
step: 30, loss: 0.0014734533615410328
step: 40, loss: 0.006787394173443317
step: 50, loss: 0.0009193720761686563
step: 60, loss: 0.0019744650926440954
step: 70, loss: 0.001909014885313809
step: 80, loss: 0.0253570806235075
step: 90, loss: 0.045144639909267426

```

```

step: 100, loss: 0.000397503754356876
step: 110, loss: 0.04591011255979538
step: 120, loss: 0.02525591105222702
step: 130, loss: 0.0056300382129848
step: 140, loss: 0.08927767723798752
step: 150, loss: 0.017303461208939552
step: 160, loss: 0.06010754778981209
step: 170, loss: 0.030545223504304886
step: 180, loss: 0.0033828658051788807
step: 190, loss: 0.020684154704213142
step: 200, loss: 0.14879456162452698
step: 210, loss: 0.0040958113968372345
step: 220, loss: 0.005162620451301336
step: 230, loss: 0.00996147096157074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887892376681614, 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: 0.002097065793350339
step: 10, loss: 0.011362612247467041
step: 20, loss: 0.002638099482282996
step: 30, loss: 0.0008520493283867836
step: 40, loss: 0.0005083063733763993
step: 50, loss: 0.0010841891635209322
step: 60, loss: 0.009406808763742447
step: 70, loss: 0.0033048910554498434
step: 80, loss: 0.001934692612849176
step: 90, loss: 0.003365958109498024
step: 100, loss: 0.005016183480620384
step: 110, loss: 0.0125547731295228
step: 120, loss: 0.0021253428421914577
step: 130, loss: 0.0014497041702270508
step: 140, loss: 0.017377736046910286
step: 150, loss: 0.0016359015135094523
step: 160, loss: 0.005137543193995953
step: 170, loss: 0.0005750560667365789
step: 180, loss: 0.004481646232306957
step: 190, loss: 0.000363118335371837
step: 200, loss: 0.006982982624322176
step: 210, loss: 0.0013391831889748573

```

```

step: 220, loss: 0.02661122940480709
step: 230, loss: 0.003904881188645959
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.987598647125141,
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: 0.009939017705619335
step: 10, loss: 0.0012901787413284183
step: 20, loss: 0.003179608378559351
step: 30, loss: 0.0009612789144739509
step: 40, loss: 0.003866110695526004
step: 50, loss: 0.001911164028570056
step: 60, loss: 0.00089325598673895
step: 70, loss: 0.00081369758117944
step: 80, loss: 0.0009424533927813172
step: 90, loss: 0.0017639750149101019
step: 100, loss: 0.0035256901755928993
step: 110, loss: 0.0010560165392234921
step: 120, loss: 0.0009137073066085577
step: 130, loss: 0.004562385380268097
step: 140, loss: 0.0003913016407750547
step: 150, loss: 0.0358000174164772
step: 160, loss: 0.0004916333127766848
step: 170, loss: 0.020525867119431496
step: 180, loss: 0.0005214893026277423
step: 190, loss: 0.0006687234272249043
step: 200, loss: 0.00039127585478127003
step: 210, loss: 0.0894169807434082
step: 220, loss: 0.00030757213244214654
step: 230, loss: 0.0008944208384491503
/usr/local/lib/python3.7/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=0.9843400447427293, f1=0.978865406006674,
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: 0.005870667286217213
step: 10, loss: 0.007749693468213081
step: 20, loss: 0.001820818753913045
step: 30, loss: 0.0009675953770056367
step: 40, loss: 0.0027165410574525595
step: 50, loss: 0.003432591911405325

```

```

step: 60, loss: 0.0006794885848648846
step: 70, loss: 0.00034601613879203796
step: 80, loss: 0.03366189822554588
step: 90, loss: 0.002779319416731596
step: 100, loss: 0.0004904197994619608
step: 110, loss: 0.01410544291138649
step: 120, loss: 0.029457291588187218
step: 130, loss: 0.0007295217947103083
step: 140, loss: 0.0005736316670663655
step: 150, loss: 0.04085591807961464
step: 160, loss: 0.017355602234601974
step: 170, loss: 0.09608807414770126
step: 180, loss: 0.0007760972948744893
step: 190, loss: 0.0014197430573403835
step: 200, loss: 0.016928106546401978
step: 210, loss: 0.00574982026591897
step: 220, loss: 0.0009611360728740692
step: 230, loss: 0.000676620053127408
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9821428571428571,
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: 0.0003916827554348856
step: 10, loss: 0.0006232610903680325
step: 20, loss: 0.00032789798569865525
step: 30, loss: 0.00030138788861222565
step: 40, loss: 0.00022340496070683002
step: 50, loss: 0.0018965299241244793
step: 60, loss: 0.00115133635699749
step: 70, loss: 0.04241817444562912
step: 80, loss: 0.0004625098081305623
step: 90, loss: 0.021484248340129852
step: 100, loss: 0.0006838358240202069
step: 110, loss: 0.0005913063650950789
step: 120, loss: 0.007588914595544338
step: 130, loss: 0.0010434980504214764
step: 140, loss: 0.0006778504466637969
step: 150, loss: 0.001448752242140472
step: 160, loss: 0.0017855934565886855
step: 170, loss: 0.0006351299816742539
step: 180, loss: 0.0005479294341057539
step: 190, loss: 0.00016115298785734922
step: 200, loss: 0.0002491832128725946
step: 210, loss: 0.004214686807245016

```

```

step: 220, loss: 0.0003805299929808825
step: 230, loss: 0.00023782291100360453
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9864559819413092,
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: 0.00038579668034799397
step: 10, loss: 0.00045957168913446367
step: 20, loss: 0.0003673759929370135
step: 30, loss: 0.00021108087094035
step: 40, loss: 0.001405875664204359
step: 50, loss: 0.0002762764343060553
step: 60, loss: 0.0010898851323872805
step: 70, loss: 0.013478667475283146
step: 80, loss: 0.0010926120448857546
step: 90, loss: 0.0004077372723259032
step: 100, loss: 0.0002763225056696683
step: 110, loss: 0.011329395696520805
step: 120, loss: 0.0005162183078937232
step: 130, loss: 0.0017602540319785476
step: 140, loss: 0.000525005511008203
step: 150, loss: 0.035295311361551285
step: 160, loss: 0.00015136392903514206
step: 170, loss: 0.00021126255160197616
step: 180, loss: 0.0006990727270022035
step: 190, loss: 0.0003620207717176527
step: 200, loss: 0.0009202365763485432
step: 210, loss: 0.001093795639462769
step: 220, loss: 0.007019623648375273
step: 230, loss: 0.0007562923710793257
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9843749999999999,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019646566943265498
step: 10, loss: 0.0004531860467977822
step: 20, loss: 0.0002882422413676977
step: 30, loss: 0.00023266897187568247
step: 40, loss: 9.47764128795825e-05
step: 50, loss: 0.0002289195981575176

```



```

step: 60, loss: 0.024686569347977638
step: 70, loss: 0.00021411773923318833
step: 80, loss: 0.010455723851919174
step: 90, loss: 0.16172270476818085
step: 100, loss: 0.0007990108570083976
step: 110, loss: 0.0016444886568933725
step: 120, loss: 0.000561414344701916
step: 130, loss: 0.0004192673950456083
step: 140, loss: 0.0017787368269637227
step: 150, loss: 0.000621231272816658
step: 160, loss: 0.002357185585424304
step: 170, loss: 0.0009776551742106676
step: 180, loss: 0.0012429089983925223
step: 190, loss: 0.0005383066600188613
step: 200, loss: 0.011234004981815815
step: 210, loss: 0.0006148756947368383
step: 220, loss: 0.003260847646743059
step: 230, loss: 0.0006255318876355886
/usr/local/lib/python3.7/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=0.9865168539325843, f1=0.9842696629213483,
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: 0.0008709730464033782
step: 10, loss: 0.00032719189766794443
step: 20, loss: 0.05691900476813316
step: 30, loss: 0.032291341572999954
step: 40, loss: 0.0010055832099169493
step: 50, loss: 0.0016080053756013513
step: 60, loss: 0.0025849754456430674
step: 70, loss: 0.0004785394121427089
step: 80, loss: 0.00018593313870951533
step: 90, loss: 0.0009176636813208461
step: 100, loss: 0.0002874645870178938
step: 110, loss: 0.00026063420227728784
step: 120, loss: 0.0003939716552849859
step: 130, loss: 0.0019523499067872763
step: 140, loss: 0.0013473144499585032
step: 150, loss: 0.0005724021466448903
step: 160, loss: 0.006254751235246658
step: 170, loss: 0.0037308968603610992
step: 180, loss: 0.0003917767317034304
step: 190, loss: 0.001025454606860876
step: 200, loss: 0.0003472002281341702
step: 210, loss: 0.0005158110288903117

```

```

step: 220, loss: 0.017606940120458603
step: 230, loss: 0.0008720502373762429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9843400447427293, f1=0.9799554565701558,
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: 0.000701434793882072
step: 10, loss: 0.0004186210862826556
step: 20, loss: 0.0006523177726194263
step: 30, loss: 0.0006763769197277725
step: 40, loss: 0.0008963198633864522
step: 50, loss: 0.0049469382502138615
step: 60, loss: 0.0007401625625789165
step: 70, loss: 0.0042539979331195354
step: 80, loss: 0.0004345816560089588
step: 90, loss: 0.0006378066609613597
step: 100, loss: 0.0010614031925797462
step: 110, loss: 0.004266309551894665
step: 120, loss: 0.0013865578221157193
step: 130, loss: 0.0010971418814733624
step: 140, loss: 0.0005031466716900468
step: 150, loss: 0.000245417671976611
step: 160, loss: 0.01782395876944065
step: 170, loss: 0.0007763077737763524
step: 180, loss: 0.037020500749349594
step: 190, loss: 0.0006963036139495671
step: 200, loss: 0.00011961498239543289
step: 210, loss: 0.0018034297972917557
step: 220, loss: 0.0006948407390154898
step: 230, loss: 0.0011198497377336025
/usr/local/lib/python3.7/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=0.9841628959276018, f1=0.9841269841269841,
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: 0.0004209191829431802
step: 10, loss: 0.0003107877855654806
step: 20, loss: 0.0005365196266211569
step: 30, loss: 0.0008255590219050646
step: 40, loss: 0.0007253524381667376
step: 50, loss: 0.0003061715979129076

```

```

step: 60, loss: 0.0005073355278000236
step: 70, loss: 0.0003672015736810863
step: 80, loss: 0.00026968243764713407
step: 90, loss: 0.0007956717163324356
step: 100, loss: 0.00046911006211303174
step: 110, loss: 0.0006662526866421103
step: 120, loss: 0.00010424415813758969
step: 130, loss: 0.003038732334971428
step: 140, loss: 0.0003161401255056262
step: 150, loss: 0.0012472722446545959
step: 160, loss: 0.0004336550773587078
step: 170, loss: 0.0004572780744638294
step: 180, loss: 0.00036981733865104616
step: 190, loss: 0.0004762596799992025
step: 200, loss: 0.0007330402149818838
step: 210, loss: 0.0003146314120385796
step: 220, loss: 0.00039002197445370257
step: 230, loss: 0.00020191411022096872
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9864559819413092,
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: 0.009853074327111244
step: 10, loss: 0.00035203713923692703
step: 20, loss: 0.0006280466332100332
step: 30, loss: 0.0004113032773602754
step: 40, loss: 0.00018311066378373653
step: 50, loss: 0.0002810250734910369
step: 60, loss: 0.04588646814227104
step: 70, loss: 0.0003477951686363667
step: 80, loss: 0.0002642220351845026
step: 90, loss: 0.0002297472528880462
step: 100, loss: 0.00018833251670002937
step: 110, loss: 0.0021046858746558428
step: 120, loss: 0.010620050132274628
step: 130, loss: 0.0002602076274342835
step: 140, loss: 0.022067654877901077
step: 150, loss: 0.0004991025198251009
step: 160, loss: 0.0020684755872935057
step: 170, loss: 0.0001805843785405159
step: 180, loss: 0.00037725025322288275
step: 190, loss: 0.0006360136321745813
step: 200, loss: 0.003776922356337309
step: 210, loss: 0.0580689013004303

```

```

step: 220, loss: 0.00033917438122443855
step: 230, loss: 0.0006079875165596604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9842696629213483, f1=0.9842696629213483,
best_f1=0.9843400447427293
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:13, 181.05it/s]
load_f1 = 0.9898989898989898
real_f1 = 0.9876819708846584
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, 169.31it/s]

```

2.1.3 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

enabled	: True
opt_level	: 02
cast_model_type	: torch.float16

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6410531997680664
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: 0.4776136577129364
step: 20, loss: 0.26390767097473145
step: 30, loss: 0.29800736904144287
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.39198365807533264
step: 50, loss: 0.6672455668449402
step: 60, loss: 0.2911683917045593
step: 70, loss: 0.2229628562927246
step: 80, loss: 0.2172715961933136
step: 90, loss: 0.16999918222427368
step: 100, loss: 0.2055845409631729
step: 110, loss: 0.1924009472131729
step: 120, loss: 0.21217626333236694
step: 130, loss: 0.15577779710292816
step: 140, loss: 0.25622624158859253
step: 150, loss: 0.11807439476251602
step: 160, loss: 0.4009971022605896
step: 170, loss: 0.12830616533756256
step: 180, loss: 0.1280282437801361
step: 190, loss: 0.13601218163967133
step: 200, loss: 0.06224867329001427
step: 210, loss: 0.0461321659386158
step: 220, loss: 0.17034733295440674
step: 230, loss: 0.23639024794101715
step: 240, loss: 0.055169977247714996
step: 250, loss: 0.02804027684032917
step: 260, loss: 0.22923457622528076
step: 270, loss: 0.2821701467037201
step: 280, loss: 0.08747202157974243
step: 290, loss: 0.11123954504728317
step: 300, loss: 0.12010218948125839
step: 310, loss: 0.19079291820526123
step: 320, loss: 0.08394522219896317
step: 330, loss: 0.07274119555950165
step: 340, loss: 0.3063580393791199

```

```

step: 350, loss: 0.16999219357967377
step: 360, loss: 0.06835293024778366
step: 370, loss: 0.05375988036394119
step: 380, loss: 0.1635858714580536
step: 390, loss: 0.01929938793182373
step: 400, loss: 0.03317749500274658
step: 410, loss: 0.2749432921409607
step: 420, loss: 0.0164693184196949
step: 430, loss: 0.10367023199796677
step: 440, loss: 0.019558696076273918
step: 450, loss: 0.020161224529147148
step: 460, loss: 0.0588124580681324
step: 470, loss: 0.030861668288707733
step: 480, loss: 0.06872591376304626
step: 490, loss: 0.20415374636650085
step: 500, loss: 0.08535058796405792
step: 510, loss: 0.10379473865032196
step: 520, loss: 0.29112255573272705
step: 530, loss: 0.11866450309753418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9235697940503432, f1=0.9297052154195011,
best_f1=0.9297052154195011
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08339633792638779
step: 10, loss: 0.05968375876545906
step: 20, loss: 0.10230959951877594
step: 30, loss: 0.13660265505313873
step: 40, loss: 0.07264135032892227
step: 50, loss: 0.04662466421723366
step: 60, loss: 0.02835194021463394
step: 70, loss: 0.014513533562421799
step: 80, loss: 0.09690044075250626
step: 90, loss: 0.03962469473481178
step: 100, loss: 0.23560646176338196
step: 110, loss: 0.010598687455058098
step: 120, loss: 0.12671972811222076
step: 130, loss: 0.007932194508612156
step: 140, loss: 0.08065558224916458
step: 150, loss: 0.03310929983854294
step: 160, loss: 0.05847505107522011

```

```

step: 170, loss: 0.02529807947576046
step: 180, loss: 0.019852224737405777
step: 190, loss: 0.03477676212787628
step: 200, loss: 0.2749656140804291
step: 210, loss: 0.05413079634308815
step: 220, loss: 0.001040064264088869
step: 230, loss: 0.05928930640220642
step: 240, loss: 0.10719120502471924
step: 250, loss: 0.017486874014139175
step: 260, loss: 0.06602494418621063
step: 270, loss: 0.10236301273107529
step: 280, loss: 0.1085863932967186
step: 290, loss: 0.03261319175362587
step: 300, loss: 0.046039335429668427
step: 310, loss: 0.016269929707050323
step: 320, loss: 0.12079708278179169
step: 330, loss: 0.11168783158063889
step: 340, loss: 0.07715918868780136
step: 350, loss: 0.012792833149433136
step: 360, loss: 0.1990956813097
step: 370, loss: 0.05133889243006706
step: 380, loss: 0.1117672398686409
step: 390, loss: 0.009233994409441948
step: 400, loss: 0.016474947333335876
step: 410, loss: 0.06790303438901901
step: 420, loss: 0.13102051615715027
step: 430, loss: 0.1559874713420868
step: 440, loss: 0.007072414271533489
step: 450, loss: 0.022861870005726814
step: 460, loss: 0.0318131186068058
step: 470, loss: 0.010816162452101707
step: 480, loss: 0.010625351220369339
step: 490, loss: 0.03774571418762207
step: 500, loss: 0.008836745284497738
step: 510, loss: 0.03440868854522705
step: 520, loss: 0.33735471963882446
step: 530, loss: 0.08717866241931915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9462465245597776, f1=0.9353187529083293,
best_f1=0.9353187529083293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09209319204092026
step: 10, loss: 0.017274197190999985
step: 20, loss: 0.01068224385380745
step: 30, loss: 0.06235218793153763
step: 40, loss: 0.18094991147518158
step: 50, loss: 0.02022162638604641
step: 60, loss: 0.04241672530770302
step: 70, loss: 0.007163624744862318
step: 80, loss: 0.021646317094564438
step: 90, loss: 0.015574599616229534
step: 100, loss: 0.060107726603746414
step: 110, loss: 0.06294839084148407
step: 120, loss: 0.0859876498579979
step: 130, loss: 0.06020328029990196
step: 140, loss: 0.03717082366347313
step: 150, loss: 0.04842095077037811
step: 160, loss: 0.01715218462049961
step: 170, loss: 0.01352704782038927
step: 180, loss: 0.019000427797436714
step: 190, loss: 0.004521268419921398
step: 200, loss: 0.016655217856168747
step: 210, loss: 0.023529469966888428
step: 220, loss: 0.12713059782981873
step: 230, loss: 0.02072206884622574
step: 240, loss: 0.028373094275593758
step: 250, loss: 0.025600697845220566
step: 260, loss: 0.026022138074040413
step: 270, loss: 0.006161956116557121
step: 280, loss: 0.012848054990172386
step: 290, loss: 0.02092590741813183
step: 300, loss: 0.0907859206199646
step: 310, loss: 0.09211745113134384
step: 320, loss: 0.10318563878536224
step: 330, loss: 0.027570553123950958
step: 340, loss: 0.013185176998376846
step: 350, loss: 0.2533947825431824
step: 360, loss: 0.008996625430881977
step: 370, loss: 0.05375028774142265
step: 380, loss: 0.004175447393208742
step: 390, loss: 0.008263993076980114
step: 400, loss: 0.1646547019481659
step: 410, loss: 0.06676676869392395
step: 420, loss: 0.00578319514170289
step: 430, loss: 0.024474119767546654
step: 440, loss: 0.19315940141677856
step: 450, loss: 0.03740796819329262
step: 460, loss: 0.06500599533319473
```

```

step: 470, loss: 0.0501403771340847
step: 480, loss: 0.12704530358314514
step: 490, loss: 0.007535808254033327
step: 500, loss: 0.031483087688684464
step: 510, loss: 0.05542471259832382
step: 520, loss: 0.005671173334121704
step: 530, loss: 0.004341010935604572
/usr/local/lib/python3.7/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=0.9391385767790261, f1=0.9302325581395349,
best_f1=0.9353187529083293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03144469112157822
step: 10, loss: 0.014738958328962326
step: 20, loss: 0.015202676877379417
step: 30, loss: 0.08687100559473038
step: 40, loss: 0.12113774567842484
step: 50, loss: 0.019389964640140533
step: 60, loss: 0.09045295417308807
step: 70, loss: 0.011065327562391758
step: 80, loss: 0.014755699783563614
step: 90, loss: 0.03244214132428169
step: 100, loss: 0.0018442010041326284
step: 110, loss: 0.13697443902492523
step: 120, loss: 0.012218422256410122
step: 130, loss: 0.04366119205951691
step: 140, loss: 0.07468177378177643
step: 150, loss: 0.053900569677352905
step: 160, loss: 0.023902764543890953
step: 170, loss: 0.015669802203774452
step: 180, loss: 0.07956714928150177
step: 190, loss: 0.04144616797566414
step: 200, loss: 0.1299707144498825
step: 210, loss: 0.007065576501190662
step: 220, loss: 0.005938849877566099
step: 230, loss: 0.020952114835381508
step: 240, loss: 0.024373751133680344
step: 250, loss: 0.15107274055480957
step: 260, loss: 0.002606517868116498
step: 270, loss: 0.016365524381399155
step: 280, loss: 0.012526712380349636
step: 290, loss: 0.008402934297919273
step: 300, loss: 0.006442543584853411
step: 310, loss: 0.008234122768044472
step: 320, loss: 0.07863326370716095

```

```

step: 330, loss: 0.0028229288291186094
step: 340, loss: 0.021118782460689545
step: 350, loss: 0.0640597864985466
step: 360, loss: 0.10615239292383194
step: 370, loss: 0.019666511565446854
step: 380, loss: 0.0038948776200413704
step: 390, loss: 0.0004325365589465946
step: 400, loss: 0.04218219220638275
step: 410, loss: 0.09551180899143219
step: 420, loss: 0.02762690745294094
step: 430, loss: 0.004266723524779081
step: 440, loss: 0.005499198101460934
step: 450, loss: 0.03278771787881851
step: 460, loss: 0.07155811041593552
step: 470, loss: 0.007631118409335613
step: 480, loss: 0.007592648733407259
step: 490, loss: 0.0041082194074988365
step: 500, loss: 0.09917983412742615
step: 510, loss: 0.14943791925907135
step: 520, loss: 0.01117826346307993
step: 530, loss: 0.04531993716955185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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=0.9535315985130112, f1=0.946931241347485,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013292822986841202
step: 10, loss: 0.004019508603960276
step: 20, loss: 0.020382627844810486
step: 30, loss: 0.0010612716432660818
step: 40, loss: 0.00033001016709022224
step: 50, loss: 0.006524748634546995
step: 60, loss: 0.07414278388023376
step: 70, loss: 0.001719295629300177
step: 80, loss: 0.0013729048660025
step: 90, loss: 0.07866004854440689
step: 100, loss: 0.01667735166847706
step: 110, loss: 0.002428903942927718
step: 120, loss: 0.11834133416414261
step: 130, loss: 0.017191771417856216
step: 140, loss: 0.05734727904200554

```

```

step: 150, loss: 0.006868541706353426
step: 160, loss: 0.0508195199072361
step: 170, loss: 0.12445145845413208
step: 180, loss: 0.0012108412338420749
step: 190, loss: 0.0017741962801665068
step: 200, loss: 0.03840727359056473
step: 210, loss: 0.008858778513967991
step: 220, loss: 0.001235826057381928
step: 230, loss: 0.0003294306807219982
step: 240, loss: 0.00180122721940279
step: 250, loss: 0.12311425060033798
step: 260, loss: 0.004449455067515373
step: 270, loss: 0.002796979621052742
step: 280, loss: 0.002003470668569207
step: 290, loss: 0.005191014613956213
step: 300, loss: 0.04837912693619728
step: 310, loss: 0.05785493925213814
step: 320, loss: 0.04632418230175972
step: 330, loss: 0.0034395663533359766
step: 340, loss: 0.01834956742823124
step: 350, loss: 0.0019098955672234297
step: 360, loss: 0.0007922331569716334
step: 370, loss: 0.0005338834598660469
step: 380, loss: 0.0005993717932142317
step: 390, loss: 0.0015810249606147408
step: 400, loss: 0.009507509879767895
step: 410, loss: 0.07603033632040024
step: 420, loss: 0.18041299283504486
step: 430, loss: 0.01426493190228939
step: 440, loss: 0.0006612301804125309
step: 450, loss: 0.04647579789161682
step: 460, loss: 0.006308317184448242
step: 470, loss: 0.015966424718499184
step: 480, loss: 0.015175262466073036
step: 490, loss: 0.007749305572360754
step: 500, loss: 0.038333021104335785
step: 510, loss: 0.000994976027868688
step: 520, loss: 0.06855059415102005
step: 530, loss: 0.010545575991272926
/usr/local/lib/python3.7/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=0.9514837494112105, f1=0.9418989135569201,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005135306157171726

```

step: 10, loss: 0.0056302654556930065
step: 20, loss: 0.0005994649836793542
step: 30, loss: 0.007434887811541557
step: 40, loss: 0.0012814842630177736
step: 50, loss: 0.0017401383956894279
step: 60, loss: 0.016465837135910988
step: 70, loss: 0.00022496251040138304
step: 80, loss: 0.000768747937399894
step: 90, loss: 0.001991430763155222
step: 100, loss: 0.02064567618072033
step: 110, loss: 0.0005356360925361514
step: 120, loss: 0.03412289172410965
step: 130, loss: 0.004415005445480347
step: 140, loss: 0.002313935663551092
step: 150, loss: 0.0033353015314787626
step: 160, loss: 0.023719586431980133
step: 170, loss: 0.012046893127262592
step: 180, loss: 0.0006750017055310309
step: 190, loss: 0.23575294017791748
step: 200, loss: 0.009136577136814594
step: 210, loss: 0.0036794908810406923
step: 220, loss: 0.012153894640505314
step: 230, loss: 0.01661282777786255
step: 240, loss: 0.010497933253645897
step: 250, loss: 0.01566224731504917
step: 260, loss: 0.0005125055904500186
step: 270, loss: 0.0004848510434385389
step: 280, loss: 0.010877853259444237
step: 290, loss: 0.0005404396797530353
step: 300, loss: 0.0022665916476398706
step: 310, loss: 0.01813535764813423
step: 320, loss: 0.00350522855296731
step: 330, loss: 0.0012603738578036427
step: 340, loss: 0.0019954333547502756
step: 350, loss: 0.0022316258400678635
step: 360, loss: 0.07763303816318512
step: 370, loss: 0.0023217611014842987
step: 380, loss: 0.0004379400343168527
step: 390, loss: 0.0007171115721575916
step: 400, loss: 0.012197955511510372
step: 410, loss: 0.002341781510040164
step: 420, loss: 0.06434961408376694
step: 430, loss: 0.00015295537014026195
step: 440, loss: 0.00016237373347394168
step: 450, loss: 0.24293455481529236
step: 460, loss: 0.003432776080444455
step: 470, loss: 0.00219158036634326
step: 480, loss: 0.005284764338284731

```

step: 490, loss: 0.002762220334261656
step: 500, loss: 0.0018490712391212583
step: 510, loss: 0.21303203701972961
step: 520, loss: 0.013591142371296883
step: 530, loss: 0.007676883600652218
/usr/local/lib/python3.7/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=0.9509116409537166, f1=0.9412861136999067,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019901860505342484
step: 10, loss: 0.006411497015506029
step: 20, loss: 0.005853429902344942
step: 30, loss: 0.011592775583267212
step: 40, loss: 0.019356105476617813
step: 50, loss: 0.004209741950035095
step: 60, loss: 0.005332069005817175
step: 70, loss: 0.032540738582611084
step: 80, loss: 0.04464134946465492
step: 90, loss: 0.0004333044635131955
step: 100, loss: 0.003059792798012495
step: 110, loss: 0.001482343184761703
step: 120, loss: 0.00957420188933611
step: 130, loss: 0.0005752926808781922
step: 140, loss: 0.002906227018684149
step: 150, loss: 0.0006406580796465278
step: 160, loss: 0.000488532125018537
step: 170, loss: 0.017994891852140427
step: 180, loss: 0.15078668296337128
step: 190, loss: 0.01471562497317791
step: 200, loss: 0.0005837198114022613
step: 210, loss: 0.01808890886604786
step: 220, loss: 0.002081881044432521
step: 230, loss: 0.00048213565605692565
step: 240, loss: 0.001974639017134905
step: 250, loss: 0.0036655976437032223
step: 260, loss: 0.002360182348638773
step: 270, loss: 0.0012268224963918328
step: 280, loss: 0.042175132781267166
step: 290, loss: 0.0016237694071605802
step: 300, loss: 0.0015770068857818842
step: 310, loss: 0.0014791139401495457
step: 320, loss: 0.0011369731510058045
step: 330, loss: 0.000554859871044755
step: 340, loss: 0.0010538449278101325

```

```

step: 350, loss: 0.0008231992833316326
step: 360, loss: 0.0032341359183192253
step: 370, loss: 0.021210666745901108
step: 380, loss: 0.01747935079038143
step: 390, loss: 0.0030093032401055098
step: 400, loss: 0.009743595495820045
step: 410, loss: 0.000273136276518926
step: 420, loss: 0.04291822761297226
step: 430, loss: 0.11012278497219086
step: 440, loss: 0.001818464370444417
step: 450, loss: 0.017149023711681366
step: 460, loss: 0.0003973868733737618
step: 470, loss: 0.11629337817430496
step: 480, loss: 0.0027118006255477667
step: 490, loss: 0.003441434819251299
step: 500, loss: 0.0021513598039746284
step: 510, loss: 0.004330481868237257
step: 520, loss: 0.01386323757469654
step: 530, loss: 0.010337181389331818
/usr/local/lib/python3.7/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=0.9371752479924421, f1=0.932573599240266,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006824991432949901
step: 10, loss: 0.007453846279531717
step: 20, loss: 0.07877836376428604
step: 30, loss: 0.0001848009560490027
step: 40, loss: 0.0001575796923134476
step: 50, loss: 0.002170566702261567
step: 60, loss: 0.000535998959094286
step: 70, loss: 0.0016256365925073624
step: 80, loss: 0.015002897009253502
step: 90, loss: 0.0003653528983704746
step: 100, loss: 0.0004135190974920988
step: 110, loss: 0.0003920221934095025
step: 120, loss: 0.01179435197263956
step: 130, loss: 0.005441519897431135
step: 140, loss: 0.0012088791700080037
step: 150, loss: 0.00019241489644628018
step: 160, loss: 0.0008812538580968976
step: 170, loss: 0.012838183902204037
step: 180, loss: 0.0007953245658427477
step: 190, loss: 0.06018028408288956
step: 200, loss: 0.004343485459685326

```

```

step: 210, loss: 0.20456616580486298
step: 220, loss: 0.02343924157321453
step: 230, loss: 0.020892184227705002
step: 240, loss: 0.0016280650161206722
step: 250, loss: 0.00027184057398699224
step: 260, loss: 0.00011797421757364646
step: 270, loss: 0.0025547214318066835
step: 280, loss: 0.00013774952094536275
step: 290, loss: 0.002493631560355425
step: 300, loss: 4.4885931856697425e-05
step: 310, loss: 6.46000771666877e-05
step: 320, loss: 0.00015001444262452424
step: 330, loss: 8.7968677689787e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 340, loss: 0.0017271587857976556
step: 350, loss: 5.1775445172097534e-05
step: 360, loss: 0.00023397790209855884
step: 370, loss: 0.0623953640460968
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 380, loss: 8.175874245353043e-05
step: 390, loss: 0.024227173998951912
step: 400, loss: 0.0043763574212789536
step: 410, loss: 0.0002625702181831002
step: 420, loss: 0.00564921647310257
step: 430, loss: 0.1347322016954422
step: 440, loss: 0.021653473377227783
step: 450, loss: 0.002104413229972124
step: 460, loss: 0.00561919528990984
step: 470, loss: 0.03726893663406372
step: 480, loss: 0.0075118825770914555
step: 490, loss: 0.005678033456206322
step: 500, loss: 0.0005349920829758048
step: 510, loss: 0.010972846299409866
step: 520, loss: 0.0005169669166207314
step: 530, loss: 0.00023228759528137743
/usr/local/lib/python3.7/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=0.9508041627246925, f1=0.9373219373219372,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001918480993481353
step: 10, loss: 0.000995173817500472
step: 20, loss: 0.0002767806581687182
step: 30, loss: 0.14297422766685486
step: 40, loss: 0.007656001951545477

```


step: 50, loss: 0.0008385457331314683
step: 60, loss: 0.0037914463318884373
step: 70, loss: 0.004920303355902433
step: 80, loss: 0.0054296753369271755
step: 90, loss: 0.06724919378757477
step: 100, loss: 0.000635709788184613
step: 110, loss: 0.01710263453423977
step: 120, loss: 0.0003052288666367531
step: 130, loss: 0.0006049457006156445
step: 140, loss: 0.004862209316343069
step: 150, loss: 0.00020409298304002732
step: 160, loss: 0.0013285847380757332
step: 170, loss: 0.002484863856807351
step: 180, loss: 0.0033367841970175505
step: 190, loss: 0.00033461919520050287
step: 200, loss: 0.00012142085324740037
step: 210, loss: 0.0001404979993822053
step: 220, loss: 0.00025002207257784903
step: 230, loss: 0.000205142394406721
step: 240, loss: 0.00017518512322567403
step: 250, loss: 0.000297560851322487
step: 260, loss: 0.00025696822558529675
step: 270, loss: 0.031059565022587776
step: 280, loss: 0.027783995494246483
step: 290, loss: 0.0004545636475086212
step: 300, loss: 0.061028871685266495
step: 310, loss: 0.027928227558732033
step: 320, loss: 0.0002544069429859519
step: 330, loss: 0.0007104690303094685
step: 340, loss: 0.003733817022293806
step: 350, loss: 0.08402007073163986
step: 360, loss: 0.15976877510547638
step: 370, loss: 9.189817501464859e-05
step: 380, loss: 0.019770674407482147
step: 390, loss: 0.002348288195207715
step: 400, loss: 0.002145562320947647
step: 410, loss: 0.0005337715265341103
step: 420, loss: 0.00044772092951461673
step: 430, loss: 0.007099812850356102
step: 440, loss: 0.00026378934853710234
step: 450, loss: 0.005567966960370541
step: 460, loss: 0.004972935654222965
step: 470, loss: 0.00013644766295328736
step: 480, loss: 0.002339705592021346
step: 490, loss: 0.014126960188150406
step: 500, loss: 0.0013589112786576152
step: 510, loss: 0.009225253947079182
step: 520, loss: 0.03765714541077614

```
step: 530, loss: 0.11144043505191803
/usr/local/lib/python3.7/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=0.9512195121951219, f1=0.9464788732394367,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004117493983358145
step: 10, loss: 0.0018803507555276155
step: 20, loss: 0.00022740697022527456
step: 30, loss: 0.0007961642113514245
step: 40, loss: 0.0011548396432772279
step: 50, loss: 0.00042967492481693625
step: 60, loss: 0.007263274397701025
step: 70, loss: 0.00010462365025887266
step: 80, loss: 0.00013034055882599205
step: 90, loss: 0.03851543739438057
step: 100, loss: 0.0004788614169228822
step: 110, loss: 0.011787819676101208
step: 120, loss: 0.00040586222894489765
step: 130, loss: 0.0019350942457094789
step: 140, loss: 0.0005238619050942361
step: 150, loss: 0.00020911124011036009
step: 160, loss: 0.034128475934267044
step: 170, loss: 9.948184015229344e-05
step: 180, loss: 0.0026419467758387327
step: 190, loss: 8.844317198963836e-05
step: 200, loss: 0.00024420255795121193
step: 210, loss: 0.00636669946834445
step: 220, loss: 0.001049902057275176
step: 230, loss: 0.00023422703088726848
step: 240, loss: 0.0007958869100548327
step: 250, loss: 0.01701231487095356
step: 260, loss: 0.0014821621589362621
step: 270, loss: 0.006672195624560118
step: 280, loss: 0.02763824537396431
step: 290, loss: 0.0004873854049947113
step: 300, loss: 0.0002098380064126104
step: 310, loss: 0.0038442290388047695
step: 320, loss: 0.0019788166973739862
step: 330, loss: 0.0007960383663885295
step: 340, loss: 6.181655044201761e-05
step: 350, loss: 0.0016919933259487152
step: 360, loss: 9.742812108015642e-05
step: 370, loss: 0.019792301580309868
step: 380, loss: 0.170232892036438
```

```

step: 390, loss: 0.0002900479012168944
step: 400, loss: 0.0004987327847629786
step: 410, loss: 0.00014941388508304954
step: 420, loss: 0.00022099337365943938
step: 430, loss: 6.799896800657734e-05
step: 440, loss: 9.26678185351193e-05
step: 450, loss: 0.1457778513431549
step: 460, loss: 0.0005431086174212396
step: 470, loss: 0.019435562193393707
step: 480, loss: 0.0012884734896942973
step: 490, loss: 0.0110927177593112
step: 500, loss: 0.001992122968658805
step: 510, loss: 0.0002590535150375217
step: 520, loss: 0.018251169472932816
step: 530, loss: 0.002539378125220537
/usr/local/lib/python3.7/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=0.9521126760563381, f1=0.9442622950819672,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001683546812273562
step: 10, loss: 0.00037622166564688087
step: 20, loss: 0.0023055693600326777
step: 30, loss: 0.00042757581104524434
step: 40, loss: 0.0005266951047815382
step: 50, loss: 0.00036478479159995914
step: 60, loss: 0.00024413850042037666
step: 70, loss: 0.0002278315951116383
step: 80, loss: 0.00018632160208653659
step: 90, loss: 0.0005237172008492053
step: 100, loss: 0.0021760037634521723
step: 110, loss: 0.0001298098941333592
step: 120, loss: 0.0006072514224797487
step: 130, loss: 0.00011806405382230878
step: 140, loss: 0.000765450531616807
step: 150, loss: 0.0003211452276445925
step: 160, loss: 0.0030695530585944653
step: 170, loss: 0.00024382521223742515
step: 180, loss: 0.0004463930381461978
step: 190, loss: 0.0018547597574070096
step: 200, loss: 0.00045503725414164364
step: 210, loss: 0.0018520097946748137
step: 220, loss: 0.005836819764226675
step: 230, loss: 0.0003590166161302477
step: 240, loss: 0.001680357032455504

```

```

step: 250, loss: 0.053840767592191696
step: 260, loss: 0.005625756923109293
step: 270, loss: 0.004438641015440226
step: 280, loss: 0.0056360336020588875
step: 290, loss: 0.0003139501204714179
step: 300, loss: 0.0003241421945858747
step: 310, loss: 0.00014616815315093845
step: 320, loss: 0.0008450138266198337
step: 330, loss: 0.00010781494347611442
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 340, loss: 0.0004692107322625816
step: 350, loss: 8.603402238804847e-05
step: 360, loss: 0.00019791873637586832
step: 370, loss: 0.0013208382297307253
step: 380, loss: 0.0014485373394563794
step: 390, loss: 0.0017986168386414647
step: 400, loss: 0.0001923427334986627
step: 410, loss: 0.0012156175216659904
step: 420, loss: 0.00013887867680750787
step: 430, loss: 0.000761836941819638
step: 440, loss: 0.00013491167919710279
step: 450, loss: 0.001631437917239964
step: 460, loss: 0.00023692265676800162
step: 470, loss: 0.0004324372857809067
step: 480, loss: 0.0005123370792716742
step: 490, loss: 5.677516310242936e-05
step: 500, loss: 7.319574797293171e-05
step: 510, loss: 0.0002015354548348114
step: 520, loss: 0.00024759789812378585
step: 530, loss: 0.00036870475742034614
/usr/local/lib/python3.7/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=0.9496738117427772, f1=0.9416666666666667,
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.005889008520171e-05
step: 10, loss: 0.0005214160773903131
step: 20, loss: 0.001975269988179207
step: 30, loss: 0.000102351994428318
step: 40, loss: 8.621605957159773e-05
step: 50, loss: 6.388716428773478e-05
step: 60, loss: 6.779443356208503e-05
step: 70, loss: 0.00010803664190461859
step: 80, loss: 6.717917858622968e-05
step: 90, loss: 0.00660286657512188

```

```
step: 100, loss: 0.2561822533607483
step: 110, loss: 0.00010010385449277237
step: 120, loss: 0.00021462592121679336
step: 130, loss: 0.00042976593249477446
step: 140, loss: 0.00017556396778672934
step: 150, loss: 0.00012350206088740379
step: 160, loss: 0.0001618512033019215
step: 170, loss: 0.00016209603927563876
step: 180, loss: 0.0002506253367755562
step: 190, loss: 0.0001813020498957485
step: 200, loss: 0.018594805151224136
step: 210, loss: 0.00014236883725970984
step: 220, loss: 0.00011449780140537769
step: 230, loss: 0.0002250366669613868
step: 240, loss: 0.00034851516829803586
step: 250, loss: 0.00010254004155285656
step: 260, loss: 6.15555327385664e-05
step: 270, loss: 0.00015107890067156404
step: 280, loss: 0.00023211754160001874
step: 290, loss: 0.001239077653735876
step: 300, loss: 0.001174261444248259
step: 310, loss: 0.03223929926753044
step: 320, loss: 0.015517709776759148
step: 330, loss: 0.0338471457362175
step: 340, loss: 0.00024141703033819795
step: 350, loss: 7.341952004935592e-05
step: 360, loss: 0.00024921895237639546
step: 370, loss: 0.0001694559323368594
step: 380, loss: 0.00015576438454445451
step: 390, loss: 0.0023423200473189354
step: 400, loss: 8.653195982333273e-05
step: 410, loss: 0.004325937479734421
step: 420, loss: 0.0001115535051212646
step: 430, loss: 0.0006396675016731024
step: 440, loss: 0.005087577272206545
step: 450, loss: 0.0017057337099686265
step: 460, loss: 8.754240843700245e-05
step: 470, loss: 0.010404725559055805
step: 480, loss: 0.000697667826898396
step: 490, loss: 9.039512951858342e-05
step: 500, loss: 0.0001472529984312132
step: 510, loss: 0.004551767837256193
step: 520, loss: 0.0006510468665510416
step: 530, loss: 0.000985126942396164
/usr/local/lib/python3.7/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=0.9527340129749768, f1=0.9488243430152145,
```

```
best_f1=0.946931241347485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.21905637672171e-05
step: 10, loss: 7.049249688861892e-05
step: 20, loss: 0.00017754109285306185
step: 30, loss: 7.756602281006053e-05
step: 40, loss: 0.0007191848126240075
step: 50, loss: 0.0015588193200528622
step: 60, loss: 0.00015224944218061864
step: 70, loss: 0.0004596611834131181
step: 80, loss: 0.00026777994935400784
step: 90, loss: 0.00011140805145259947
step: 100, loss: 6.1001999711152166e-05
step: 110, loss: 7.761794404359534e-05
step: 120, loss: 5.633191540255211e-05
step: 130, loss: 4.466369500732981e-05
step: 140, loss: 0.00029535588691942394
step: 150, loss: 0.0013717911206185818
step: 160, loss: 0.0007898105541244149
step: 170, loss: 2.0573481378960423e-05
step: 180, loss: 0.00016983880777843297
step: 190, loss: 5.9675759985111654e-05
step: 200, loss: 6.287715950747952e-05
step: 210, loss: 5.065560617367737e-05
step: 220, loss: 5.933025022386573e-05
step: 230, loss: 0.00029566840385086834
step: 240, loss: 0.00017544825095683336
step: 250, loss: 7.199942047009245e-05
step: 260, loss: 0.00013358907017391175
step: 270, loss: 0.0009494540863670409
step: 280, loss: 9.736057108966634e-05
step: 290, loss: 7.07576036802493e-05
step: 300, loss: 5.49878750462085e-05
step: 310, loss: 9.007701009977609e-05
step: 320, loss: 5.3633742936654016e-05
step: 330, loss: 0.00015731582243461162
step: 340, loss: 0.0011339125921949744
step: 350, loss: 4.783388430951163e-05
step: 360, loss: 0.09634476155042648
step: 370, loss: 0.00036495571839623153
step: 380, loss: 0.00016244682774413377
step: 390, loss: 0.00022246439766604453
step: 400, loss: 0.00015338334196712822
step: 410, loss: 0.00012679783685598522
step: 420, loss: 8.681986946612597e-05
step: 430, loss: 0.0003103717463091016
```

```

step: 440, loss: 5.137754851602949e-05
step: 450, loss: 0.1552739292383194
step: 460, loss: 0.00014449261652771384
step: 470, loss: 0.0012218368938192725
step: 480, loss: 3.546154039213434e-05
step: 490, loss: 0.00022348822676576674
step: 500, loss: 0.00022111760335974395
step: 510, loss: 0.0002137538103852421
step: 520, loss: 0.0009026016923598945
step: 530, loss: 9.278654033550993e-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=0.9557021677662583, f1=0.947022972339428,
best_f1=0.947022972339428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001486199558712542
step: 10, loss: 0.0014747236855328083
step: 20, loss: 0.00010483126970939338
step: 30, loss: 0.002086395863443613
step: 40, loss: 0.0002686189254745841
step: 50, loss: 0.0053170956671237946
step: 60, loss: 0.00010783245670609176
step: 70, loss: 0.00020396552281454206
step: 80, loss: 0.00023301907640416175
step: 90, loss: 0.00014527879829984158
step: 100, loss: 0.00012511180830188096
step: 110, loss: 0.0001747684000292793
step: 120, loss: 8.164123573806137e-05
step: 130, loss: 8.737476309761405e-05
step: 140, loss: 0.00023134369985200465
step: 150, loss: 0.0008367938571609557
step: 160, loss: 0.00015361575060524046
step: 170, loss: 0.00010593170736683533
step: 180, loss: 4.806220385944471e-05
step: 190, loss: 7.622939301654696e-05
step: 200, loss: 0.00011504354188218713
step: 210, loss: 0.0015388766769319773
step: 220, loss: 6.0404683608794585e-05
step: 230, loss: 0.0001119101871154271
step: 240, loss: 0.0013761507580056787
step: 250, loss: 0.0076738581992685795

```

```

step: 260, loss: 0.0003641262883320451
step: 270, loss: 0.0004154751368332654
step: 280, loss: 0.00041344837518408895
step: 290, loss: 4.685363455791958e-05
step: 300, loss: 0.00019133530440740287
step: 310, loss: 9.749074524734169e-05
step: 320, loss: 0.0004463551740627736
step: 330, loss: 0.000622231571469456
step: 340, loss: 0.00010783046309370548
step: 350, loss: 5.98948827246204e-05
step: 360, loss: 7.273277878994122e-05
step: 370, loss: 0.004743644967675209
step: 380, loss: 0.00026204853202216327
step: 390, loss: 0.0028471876867115498
step: 400, loss: 0.010798463597893715
step: 410, loss: 3.617561742430553e-05
step: 420, loss: 4.9116279114969075e-05
step: 430, loss: 0.0009381826384924352
step: 440, loss: 0.006453900132328272
step: 450, loss: 0.0009340202668681741
step: 460, loss: 0.002077521523460746
step: 470, loss: 9.535733988741413e-05
step: 480, loss: 0.00015929003711789846
step: 490, loss: 0.00030972884269431233
step: 500, loss: 0.0015257400227710605
step: 510, loss: 0.0020420453511178493
step: 520, loss: 4.6390552597586066e-05
step: 530, loss: 0.0019224146381020546
/usr/local/lib/python3.7/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=0.9536178107606679, f1=0.9497464269248502,
best_f1=0.947022972339428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011918208474526182
step: 10, loss: 0.001743054948747158
step: 20, loss: 0.0008600398432463408
step: 30, loss: 0.002416069619357586
step: 40, loss: 7.526092667831108e-05
step: 50, loss: 0.0005097496323287487
step: 60, loss: 0.0004279833228792995
step: 70, loss: 0.00018638464098330587
step: 80, loss: 5.978584522381425e-05
step: 90, loss: 5.861672980245203e-05
step: 100, loss: 2.0633198801078834e-05
step: 110, loss: 0.0003887033089995384

```



```
step: 120, loss: 4.627781891031191e-05
step: 130, loss: 0.00021683478553313762
step: 140, loss: 5.572885129367933e-05
step: 150, loss: 9.805030276766047e-05
step: 160, loss: 9.226910333381966e-05
step: 170, loss: 3.238841600250453e-05
step: 180, loss: 0.00010940904030576348
step: 190, loss: 0.0003923483891412616
step: 200, loss: 0.0011509613832458854
step: 210, loss: 4.071359217050485e-05
step: 220, loss: 8.317790343426168e-05
step: 230, loss: 0.00011136670946143568
step: 240, loss: 5.446936120279133e-05
step: 250, loss: 5.6975499319378287e-05
step: 260, loss: 4.888115290668793e-05
step: 270, loss: 9.838271944317967e-05
step: 280, loss: 0.00017268213559873402
step: 290, loss: 5.105424497742206e-05
step: 300, loss: 6.42876693746075e-05
step: 310, loss: 0.08290035277605057
step: 320, loss: 0.000240478795603849
step: 330, loss: 8.667857036925852e-05
step: 340, loss: 8.564317249692976e-05
step: 350, loss: 0.0014614008832722902
step: 360, loss: 7.591769826831296e-05
step: 370, loss: 0.0002920891856774688
step: 380, loss: 7.230541086755693e-05
step: 390, loss: 9.87105377134867e-05
step: 400, loss: 0.00017712742555886507
step: 410, loss: 9.378221147926524e-05
step: 420, loss: 9.578339086147025e-05
step: 430, loss: 3.118891254416667e-05
step: 440, loss: 0.15355022251605988
step: 450, loss: 0.003787447465583682
step: 460, loss: 0.0014857081696391106
step: 470, loss: 5.3064621170051396e-05
step: 480, loss: 0.001669514225795865
step: 490, loss: 0.0009526055655442178
step: 500, loss: 0.028679771348834038
step: 510, loss: 5.5673997849226e-05
step: 520, loss: 7.055460446281359e-05
step: 530, loss: 3.567776730051264e-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=0.9541804440245631, f1=0.9481132075471699,
best_f1=0.947022972339428
mv: cannot stat '*_dev.pt': No such file or directory
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

5742it [00:22, 258.80it/s]

load_f1 = 0.9546940681924334

real_f1 = 0.9532710280373832

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, 199.24it/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/

```

```

Downloading: 100% 481/481 [00:00<00:00, 365kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 1.73MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 1.31MB/s]
Downloading: 100% 501M/501M [00:13<00:00, 37.7MB/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: 0.5475060343742371
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4192194938659668
step: 20, loss: 0.4085792005062103
step: 30, loss: 0.4120912551879883
step: 40, loss: 0.27877315878868103
step: 50, loss: 0.41970428824424744
step: 60, loss: 0.5299904346466064
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.3458731472492218
step: 80, loss: 0.3744012713432312
step: 90, loss: 0.24934378266334534
step: 100, loss: 0.29081565141677856
step: 110, loss: 0.3148813545703888
step: 120, loss: 0.3791716694831848
step: 130, loss: 0.24214166402816772
step: 140, loss: 0.4741656184196472
step: 150, loss: 0.36170297861099243
step: 160, loss: 0.5172631144523621
step: 170, loss: 0.25953999161720276
step: 180, loss: 0.39638352394104004
step: 190, loss: 0.6977786421775818
step: 200, loss: 0.35931891202926636
step: 210, loss: 0.5044553875923157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3669426441192627
step: 10, loss: 0.16370432078838348
step: 20, loss: 0.47080937027931213
step: 30, loss: 0.5000967979431152
step: 40, loss: 0.4513693153858185
step: 50, loss: 0.32134485244750977
step: 60, loss: 0.33093634247779846
step: 70, loss: 0.3589194416999817
step: 80, loss: 0.32452288269996643
step: 90, loss: 0.20680733025074005
step: 100, loss: 0.4579811990261078
step: 110, loss: 0.41764628887176514
step: 120, loss: 0.13118450343608856
step: 130, loss: 0.24865727126598358
step: 140, loss: 0.1894187182188034
step: 150, loss: 0.3384535312652588
step: 160, loss: 0.09390611201524734
step: 170, loss: 0.37323155999183655
step: 180, loss: 0.305441677570343
step: 190, loss: 0.3252173066139221
step: 200, loss: 0.14767004549503326
step: 210, loss: 0.16888615489006042
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.537117903930131, f1=0.55, best_f1=0.55
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11848083883523941
step: 10, loss: 0.07563940435647964
step: 20, loss: 0.2848118245601654
step: 30, loss: 0.0550009161233902
step: 40, loss: 0.3412099778652191
step: 50, loss: 0.49531233310699463
step: 60, loss: 0.3219683766365051
step: 70, loss: 0.13644656538963318
step: 80, loss: 0.2576616406440735
step: 90, loss: 0.11002644151449203
step: 100, loss: 0.2379886507987976
step: 110, loss: 0.10986697673797607
step: 120, loss: 0.11080457270145416

```

```

step: 130, loss: 0.3294955790042877
step: 140, loss: 0.13280777633190155
step: 150, loss: 0.2936226725578308
step: 160, loss: 0.20672933757305145
step: 170, loss: 0.2796919643878937
step: 180, loss: 0.1381911337375641
step: 190, loss: 0.15973185002803802
step: 200, loss: 0.13680224120616913
step: 210, loss: 0.2463354468345642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6145038167938932, f1=0.6323809523809524,
best_f1=0.6323809523809524
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1050807535648346
step: 10, loss: 0.14240096509456635
step: 20, loss: 0.27712908387184143
step: 30, loss: 0.140095055103302
step: 40, loss: 0.06049547716975212
step: 50, loss: 0.1473277509212494
step: 60, loss: 0.09543930739164352
step: 70, loss: 0.08570719510316849
step: 80, loss: 0.07212355732917786
step: 90, loss: 0.03858532756567001
step: 100, loss: 0.15774621069431305
step: 110, loss: 0.38067758083343506
step: 120, loss: 0.3064025044441223
step: 130, loss: 0.38939186930656433
step: 140, loss: 0.2594287097454071
step: 150, loss: 0.14066417515277863
step: 160, loss: 0.21146202087402344
step: 170, loss: 0.10424946248531342
step: 180, loss: 0.06889484077692032
step: 190, loss: 0.13183414936065674
step: 200, loss: 0.039157260209321976
step: 210, loss: 0.43662598729133606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6352941176470588, f1=0.6506024096385543,
best_f1=0.6506024096385543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1514880359172821
step: 10, loss: 0.0368863046169281
step: 20, loss: 0.12477368116378784
step: 30, loss: 0.13918671011924744
step: 40, loss: 0.07001124322414398
step: 50, loss: 0.1265668123960495
step: 60, loss: 0.15781691670417786
step: 70, loss: 0.2097218632698059
step: 80, loss: 0.06964045763015747
step: 90, loss: 0.08147967606782913
step: 100, loss: 0.03214964643120766
step: 110, loss: 0.03358667343854904
step: 120, loss: 0.1011008620262146
step: 130, loss: 0.12304659932851791
step: 140, loss: 0.09411761909723282
step: 150, loss: 0.09220247715711594
step: 160, loss: 0.04133979231119156
step: 170, loss: 0.11267562955617905
step: 180, loss: 0.21081693470478058
step: 190, loss: 0.1671629697084427
step: 200, loss: 0.09891501814126968
step: 210, loss: 0.07011056691408157
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6415094339622641, f1=0.6388308977035491,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08642378449440002
step: 10, loss: 0.09568468481302261
step: 20, loss: 0.01536202896386385
step: 30, loss: 0.16768454015254974
step: 40, loss: 0.0978696271777153
step: 50, loss: 0.09837418794631958
step: 60, loss: 0.19898946583271027

```

```

step: 70, loss: 0.14835263788700104
step: 80, loss: 0.11966212093830109
step: 90, loss: 0.04959942400455475
step: 100, loss: 0.16231673955917358
step: 110, loss: 0.1056709736585617
step: 120, loss: 0.011260390281677246
step: 130, loss: 0.043461594730615616
step: 140, loss: 0.08612941205501556
step: 150, loss: 0.025658197700977325
step: 160, loss: 0.023125076666474342
step: 170, loss: 0.14514102041721344
step: 180, loss: 0.022906163707375526
step: 190, loss: 0.03300909698009491
step: 200, loss: 0.3254722058773041
step: 210, loss: 0.1330813318490982
/usr/local/lib/python3.7/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=0.6129753914988815, f1=0.6405228758169934,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08630023151636124
step: 10, loss: 0.015479018911719322
step: 20, loss: 0.023380370810627937
step: 30, loss: 0.08632446825504303
step: 40, loss: 0.0028717753011733294
step: 50, loss: 0.026789570227265358
step: 60, loss: 0.03489065170288086
step: 70, loss: 0.09738609194755554
step: 80, loss: 0.04438593238592148
step: 90, loss: 0.034641895443201065
step: 100, loss: 0.07006800174713135
step: 110, loss: 0.034115102142095566
step: 120, loss: 0.07424519211053848
step: 130, loss: 0.2832280695438385
step: 140, loss: 0.10854081064462662
step: 150, loss: 0.15445534884929657
step: 160, loss: 0.19733567535877228
step: 170, loss: 0.08219332993030548
step: 180, loss: 0.05779499188065529
step: 190, loss: 0.009514587931334972
step: 200, loss: 0.11713308840990067
step: 210, loss: 0.24671967327594757
/usr/local/lib/python3.7/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=0.6223021582733813, f1=0.6272401433691756,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03846431523561478
step: 10, loss: 0.017206337302923203
step: 20, loss: 0.008121682330965996
step: 30, loss: 0.025744084268808365
step: 40, loss: 0.001285431208088994
step: 50, loss: 0.006567764095962048
step: 60, loss: 0.004820812493562698
step: 70, loss: 0.17255575954914093
step: 80, loss: 0.17193365097045898
step: 90, loss: 0.035664018243551254
step: 100, loss: 0.09307321161031723
step: 110, loss: 0.04783233255147934
step: 120, loss: 0.038215164095163345
step: 130, loss: 0.00659555708989501
step: 140, loss: 0.11684661358594894
step: 150, loss: 0.13144662976264954
step: 160, loss: 0.0774916335940361
step: 170, loss: 0.03972387686371803
step: 180, loss: 0.07231878489255905
step: 190, loss: 0.05069033429026604
step: 200, loss: 0.027724670246243477
step: 210, loss: 0.05122580751776695
/usr/local/lib/python3.7/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=0.6310344827586207, f1=0.6112054329371815,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1544766128063202
step: 10, loss: 0.022094199433922768
step: 20, loss: 0.046243757009506226
step: 30, loss: 0.006351597607135773
step: 40, loss: 0.0008551526698283851
step: 50, loss: 0.12003976106643677
step: 60, loss: 0.0015548373339697719
step: 70, loss: 0.051153041422367096
step: 80, loss: 0.09023191034793854
step: 90, loss: 0.06390698254108429
step: 100, loss: 0.029284559190273285
step: 110, loss: 0.05525222420692444
step: 120, loss: 0.013723446056246758

```

```

step: 130, loss: 0.1144665852189064
step: 140, loss: 0.015274116769433022
step: 150, loss: 0.048033613711595535
step: 160, loss: 0.33123552799224854
step: 170, loss: 0.040807586163282394
step: 180, loss: 0.04859264940023422
step: 190, loss: 0.02064920775592327
step: 200, loss: 0.004408596083521843
step: 210, loss: 0.1617612987756729
/usr/local/lib/python3.7/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=0.6170678336980306, f1=0.6439909297052154,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006089388392865658
step: 10, loss: 0.007215231657028198
step: 20, loss: 0.007881268858909607
step: 30, loss: 0.0007054340676404536
step: 40, loss: 0.0035177255049347878
step: 50, loss: 0.05839994177222252
step: 60, loss: 0.00949932262301445
step: 70, loss: 0.01977979578077793
step: 80, loss: 0.024303385987877846
step: 90, loss: 0.03175632655620575
step: 100, loss: 0.16611774265766144
step: 110, loss: 0.009186332114040852
step: 120, loss: 0.049222975969314575
step: 130, loss: 0.05073372274637222
step: 140, loss: 0.0033486851025372744
step: 150, loss: 0.00962571520358324
step: 160, loss: 0.059716928750276566
step: 170, loss: 0.0506901778280735
step: 180, loss: 0.014460490085184574
step: 190, loss: 0.02819492481648922
step: 200, loss: 0.051947180181741714
step: 210, loss: 0.0016626986907795072
/usr/local/lib/python3.7/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=0.6322067594433399, f1=0.6666666666666666,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06879120320081711

```

```

step: 10, loss: 0.005334039684385061
step: 20, loss: 0.004534280858933926
step: 30, loss: 0.004460558760911226
step: 40, loss: 0.006208166480064392
step: 50, loss: 0.09847467392683029
step: 60, loss: 0.03378371149301529
step: 70, loss: 0.020532850176095963
step: 80, loss: 0.006071555428206921
step: 90, loss: 0.1929388791322708
step: 100, loss: 0.023307427763938904
step: 110, loss: 0.02897661365568638
step: 120, loss: 0.04546008259057999
step: 130, loss: 0.004255687817931175
step: 140, loss: 0.019502053037285805
step: 150, loss: 0.04282231256365776
step: 160, loss: 0.0008216026471927762
step: 170, loss: 0.0110095189884305
step: 180, loss: 0.009565891698002815
step: 190, loss: 0.03788873553276062
step: 200, loss: 0.013511833734810352
step: 210, loss: 0.03959844633936882
/usr/local/lib/python3.7/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=0.6311637080867851, f1=0.6516393442622951,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017229719087481499
step: 10, loss: 0.0151385348290205
step: 20, loss: 0.01785723865032196
step: 30, loss: 0.008707272820174694
step: 40, loss: 0.0003498812729958445
step: 50, loss: 0.0015356853837147355
step: 60, loss: 0.0037837871350347996
step: 70, loss: 0.009081773459911346
step: 80, loss: 0.0339091531932354
step: 90, loss: 0.01029975526034832
step: 100, loss: 0.0022925708908587694
step: 110, loss: 0.1370517909526825
step: 120, loss: 0.0022741775028407574
step: 130, loss: 0.03369038552045822
step: 140, loss: 0.07021086663007736
step: 150, loss: 0.000163894597790204
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.0006040356238372624
step: 170, loss: 0.1524653285741806

```

```

step: 180, loss: 0.002118697389960289
step: 190, loss: 0.03485334292054176
step: 200, loss: 0.0028631456661969423
step: 210, loss: 0.04052410647273064
/usr/local/lib/python3.7/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=0.6307385229540918, f1=0.6448979591836735,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028284990694373846
step: 10, loss: 0.025921232998371124
step: 20, loss: 0.005719548091292381
step: 30, loss: 0.014841757714748383
step: 40, loss: 0.0010234410874545574
step: 50, loss: 0.17381179332733154
step: 60, loss: 0.01583748497068882
step: 70, loss: 0.005410260055214167
step: 80, loss: 0.17559699714183807
step: 90, loss: 0.027602486312389374
step: 100, loss: 0.0020028813742101192
step: 110, loss: 0.04581542685627937
step: 120, loss: 0.0021634099539369345
step: 130, loss: 0.005050220992416143
step: 140, loss: 0.006664101034402847
step: 150, loss: 0.002333735814318061
step: 160, loss: 0.03534796088933945
step: 170, loss: 0.0013205045834183693
step: 180, loss: 0.002332722768187523
step: 190, loss: 0.02463538572192192
step: 200, loss: 0.005082267336547375
step: 210, loss: 0.14925217628479004
/usr/local/lib/python3.7/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=0.6367041198501873, f1=0.6425855513307984,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002356564626097679
step: 10, loss: 0.09091349691152573
step: 20, loss: 0.002264933427795768
step: 30, loss: 0.0014448290457949042
step: 40, loss: 0.016473736613988876
step: 50, loss: 0.0051567391492426395

```

```

step: 60, loss: 0.12036645412445068
step: 70, loss: 0.005105873569846153
step: 80, loss: 0.017929673194885254
step: 90, loss: 0.0005132087389938533
step: 100, loss: 0.014441807754337788
step: 110, loss: 0.002449811203405261
step: 120, loss: 0.001108461758121848
step: 130, loss: 0.016711294651031494
step: 140, loss: 0.0011954933870583773
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 150, loss: 0.035586025565862656
step: 160, loss: 0.013954708352684975
step: 170, loss: 0.03810713812708855
step: 180, loss: 0.0007504393579438329
step: 190, loss: 0.05636712163686752
step: 200, loss: 0.0007891888380981982
step: 210, loss: 0.0016901949420571327
/usr/local/lib/python3.7/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=0.617283950617284, f1=0.6515463917525772,
best_f1=0.6388308977035491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026621201541274786
step: 10, loss: 0.002973872469738126
step: 20, loss: 0.05286013334989548
step: 30, loss: 0.010553041473031044
step: 40, loss: 0.005619758740067482
step: 50, loss: 0.0006612472352571785
step: 60, loss: 0.013018688187003136
step: 70, loss: 0.0026698322035372257
step: 80, loss: 0.06327628344297409
step: 90, loss: 0.005336458794772625
step: 100, loss: 0.0024227285757660866
step: 110, loss: 0.004374115727841854
step: 120, loss: 0.00422790739685297
step: 130, loss: 0.001439272309653461
step: 140, loss: 0.010767695493996143
step: 150, loss: 0.02774989604949951
step: 160, loss: 0.003342424286529422
step: 170, loss: 0.055437274277210236
step: 180, loss: 0.002881844062358141
step: 190, loss: 0.0012344305869191885
step: 200, loss: 0.0008217753493227065
step: 210, loss: 0.028741173446178436
/usr/local/lib/python3.7/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=0.6086956521739131, f1=0.650103519668737,
best_f1=0.6388308977035491
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, 339.52it/s]
load_f1 = 0.6335403726708074
real_f1 = 0.6359832635983262
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:25, 171.41it/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: 0.4288548231124878
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4566175043582916
step: 20, loss: 0.26247233152389526
step: 30, loss: 0.3509836792945862
step: 40, loss: 0.26394426822662354
step: 50, loss: 0.32558155059814453
step: 60, loss: 0.44326525926589966
step: 70, loss: 0.4025137424468994
step: 80, loss: 0.17092333734035492
step: 90, loss: 0.29263654351234436
step: 100, loss: 0.4466719627380371
step: 110, loss: 0.19748635590076447
step: 120, loss: 0.32798469066619873
step: 130, loss: 0.341731995344162
step: 140, loss: 0.18872126936912537
step: 150, loss: 0.27670714259147644
step: 160, loss: 0.19449450075626373
step: 170, loss: 0.3873932361602783
step: 180, loss: 0.22969332337379456
step: 190, loss: 0.12492495775222778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4526558891454966, f1=0.5188470066518847,

```



```

best_f1=0.5188470066518847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27371078729629517
step: 10, loss: 0.2380666583776474
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.713784396648407
step: 30, loss: 0.19283238053321838
step: 40, loss: 0.23626410961151123
step: 50, loss: 0.34026387333869934
step: 60, loss: 0.2060551941394806
step: 70, loss: 0.2313794493675232
step: 80, loss: 0.06335984915494919
step: 90, loss: 0.1144426167011261
step: 100, loss: 0.11302945762872696
step: 110, loss: 0.05478953942656517
step: 120, loss: 0.15811878442764282
step: 130, loss: 0.11300838738679886
step: 140, loss: 0.19309131801128387
step: 150, loss: 0.23621486127376556
step: 160, loss: 0.08736107498407364
step: 170, loss: 0.029644062742590904
step: 180, loss: 0.06463753432035446
step: 190, loss: 0.11520742624998093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8053333333333333, f1=0.7978436657681941,
best_f1=0.7978436657681941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10430026799440384
step: 10, loss: 0.0819181427359581
step: 20, loss: 0.01186447124928236
step: 30, loss: 0.02777501568198204
step: 40, loss: 0.015544845722615719
step: 50, loss: 0.1386820375919342
step: 60, loss: 0.05362938344478607
step: 70, loss: 0.1509748101234436
step: 80, loss: 0.3041498363018036
step: 90, loss: 0.06299461424350739
step: 100, loss: 0.09403858333826065

```

```

step: 110, loss: 0.27506664395332336
step: 120, loss: 0.4146132171154022
step: 130, loss: 0.10566169023513794
step: 140, loss: 0.015924841165542603
step: 150, loss: 0.251449316740036
step: 160, loss: 0.05025660619139671
step: 170, loss: 0.18330854177474976
step: 180, loss: 0.16574588418006897
step: 190, loss: 0.026820501312613487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8440860215053764, f1=0.8412256267409471,
best_f1=0.8412256267409471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012196077965199947
step: 10, loss: 0.041403643786907196
step: 20, loss: 0.008654399774968624
step: 30, loss: 0.005533084738999605
step: 40, loss: 0.07496985793113708
step: 50, loss: 0.0618736557662487
step: 60, loss: 0.08670750260353088
step: 70, loss: 0.02920600026845932
step: 80, loss: 0.009136198088526726
step: 90, loss: 0.04643981531262398
step: 100, loss: 0.12603391706943512
step: 110, loss: 0.13550350069999695
step: 120, loss: 0.015139274299144745
step: 130, loss: 0.03908596187829971
step: 140, loss: 0.1528458148241043
step: 150, loss: 0.1456831842660904
step: 160, loss: 0.02810279093682766
step: 170, loss: 0.010540150105953217
step: 180, loss: 0.08933249115943909
step: 190, loss: 0.007950286380946636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8587570621468926, f1=0.8387096774193548,
best_f1=0.8387096774193548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06571350246667862
step: 10, loss: 0.08819311857223511
step: 20, loss: 0.1013399139046669
step: 30, loss: 0.029008865356445312
step: 40, loss: 0.05145685747265816
step: 50, loss: 0.0390777625143528
step: 60, loss: 0.0038084520492702723
step: 70, loss: 0.02018168568611145
step: 80, loss: 0.006997869350016117
step: 90, loss: 0.11177463829517365
step: 100, loss: 0.1216685026884079
step: 110, loss: 0.048943184316158295
step: 120, loss: 0.01852243021130562
step: 130, loss: 0.26327428221702576
step: 140, loss: 0.016824128106236458
step: 150, loss: 0.014861969277262688
step: 160, loss: 0.004745223093777895
step: 170, loss: 0.0016256296075880527
step: 180, loss: 0.005026835482567549
step: 190, loss: 0.002178947441279888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8726287262872628, f1=0.8435754189944134,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18163377046585083
step: 10, loss: 0.03053869865834713
step: 20, loss: 0.0023372843861579895
step: 30, loss: 0.08204340934753418
step: 40, loss: 0.02993318811058998
step: 50, loss: 0.05664385110139847
step: 60, loss: 0.021462345495820045
step: 70, loss: 0.10706944018602371
step: 80, loss: 0.04957476258277893
step: 90, loss: 0.009258344769477844
step: 100, loss: 0.09724009037017822

```

```

step: 110, loss: 0.0032394311856478453
step: 120, loss: 0.007947885431349277
step: 130, loss: 0.08003837615251541
step: 140, loss: 0.007267233915627003
step: 150, loss: 0.0032208398915827274
step: 160, loss: 0.04210694134235382
step: 170, loss: 0.10150128602981567
step: 180, loss: 0.008111303672194481
step: 190, loss: 0.003746042726561427
/usr/local/lib/python3.7/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=0.8595041322314049, f1=0.8491620111731845,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007070232182741165
step: 10, loss: 0.0457807257771492
step: 20, loss: 0.0025475851725786924
step: 30, loss: 0.00276504666544497
step: 40, loss: 0.004332822747528553
step: 50, loss: 0.0022017110604792833
step: 60, loss: 0.03666781261563301
step: 70, loss: 0.0023923651315271854
step: 80, loss: 0.004204300232231617
step: 90, loss: 0.0013648758176714182
step: 100, loss: 0.3872354328632355
step: 110, loss: 0.08815612643957138
step: 120, loss: 0.17974452674388885
step: 130, loss: 0.013019341044127941
step: 140, loss: 0.03554774075746536
step: 150, loss: 0.01948402263224125
step: 160, loss: 0.013302190229296684
step: 170, loss: 0.008031359873712063
step: 180, loss: 0.014486439526081085
step: 190, loss: 0.012200187891721725
/usr/local/lib/python3.7/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=0.8547008547008547,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003078218549489975
step: 10, loss: 0.023428218439221382
step: 20, loss: 0.002928889822214842

```

```

step: 30, loss: 0.20733323693275452
step: 40, loss: 0.012424931861460209
step: 50, loss: 0.00643658172339201
step: 60, loss: 0.030478404834866524
step: 70, loss: 0.004563593305647373
step: 80, loss: 0.014528967440128326
step: 90, loss: 0.0022047569509595633
step: 100, loss: 0.03031698428094387
step: 110, loss: 0.037798065692186356
step: 120, loss: 0.0012034053215757012
step: 130, loss: 0.04326840490102768
step: 140, loss: 0.003189951414242387
step: 150, loss: 0.010504703037440777
step: 160, loss: 0.001041662646457553
step: 170, loss: 0.001498694997280836
step: 180, loss: 0.02902979962527752
step: 190, loss: 0.002303674817085266
/usr/local/lib/python3.7/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=0.8346883468834688, f1=0.8435754189944134,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035174690186977386
step: 10, loss: 0.000307630019960925
step: 20, loss: 0.0018272591987624764
step: 30, loss: 0.0021156154107302427
step: 40, loss: 0.04111682251095772
step: 50, loss: 0.002634285483509302
step: 60, loss: 0.004903859458863735
step: 70, loss: 0.0017676305724307895
step: 80, loss: 0.003615631489083171
step: 90, loss: 0.05995938926935196
step: 100, loss: 0.0008515265653841197
step: 110, loss: 0.0012091518146917224
step: 120, loss: 0.0019227721495553851
step: 130, loss: 0.00037686442374251783
step: 140, loss: 0.07749998569488525
step: 150, loss: 0.0004198602691758424
step: 160, loss: 0.00019794091349467635
step: 170, loss: 0.0023243073374032974
step: 180, loss: 0.0029000621289014816
step: 190, loss: 0.0004041357315145433
/usr/local/lib/python3.7/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=0.8484848484848485, f1=0.8350515463917526,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000441745767602697
step: 10, loss: 0.0009777627419680357
step: 20, loss: 0.0004329691582825035
step: 30, loss: 0.002772530307993293
step: 40, loss: 0.004293725360184908
step: 50, loss: 0.000415921735111624
step: 60, loss: 0.0012058160500600934
step: 70, loss: 0.0002399195946054533
step: 80, loss: 0.0009769065072759986
step: 90, loss: 0.0001507240376668051
step: 100, loss: 0.0003100273315794766
step: 110, loss: 0.00043363080476410687
step: 120, loss: 0.003960315603762865
step: 130, loss: 0.00023523395066149533
step: 140, loss: 0.004414531402289867
step: 150, loss: 0.0002859215310309082
step: 160, loss: 0.00043288161396048963
step: 170, loss: 0.0004488361009862274
step: 180, loss: 0.0036339040379971266
step: 190, loss: 0.000283699861029163
/usr/local/lib/python3.7/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=0.8431876606683805, f1=0.8602150537634409,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028503856156021357
step: 10, loss: 0.006059486418962479
step: 20, loss: 0.00013576375204138458
step: 30, loss: 0.02117305062711239
step: 40, loss: 0.000304176879581064
step: 50, loss: 0.004768006037920713
step: 60, loss: 0.02300337329506874
step: 70, loss: 0.0002822637907229364
step: 80, loss: 0.016132649034261703
step: 90, loss: 0.0001798134035198018
step: 100, loss: 0.00047653206274844706
step: 110, loss: 0.026944240555167198
step: 120, loss: 0.0005653008120134473
step: 130, loss: 0.0013309090863913298
step: 140, loss: 0.00031751723145134747

```

```

step: 150, loss: 0.012552346102893353
step: 160, loss: 0.00032633685623295605
step: 170, loss: 0.011996421962976456
step: 180, loss: 0.0007712949300184846
step: 190, loss: 0.0001713693782221526
/usr/local/lib/python3.7/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=0.8296296296296296, f1=0.8491048593350383,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011317159980535507
step: 10, loss: 0.0019701814744621515
step: 20, loss: 0.0007854156428948045
step: 30, loss: 0.003056248649954796
step: 40, loss: 0.0012115953722968698
step: 50, loss: 0.030312126502394676
step: 60, loss: 0.00267288857139647
step: 70, loss: 0.00040095497388392687
step: 80, loss: 0.0014374115271493793
step: 90, loss: 0.00153733033221215
step: 100, loss: 0.00022790675575379282
step: 110, loss: 0.005210140720009804
step: 120, loss: 0.0007748683565296233
step: 130, loss: 0.00014939196989871562
step: 140, loss: 0.00038172377389855683
step: 150, loss: 0.0020291877444833517
step: 160, loss: 0.00015173810243140906
step: 170, loss: 0.018045969307422638
step: 180, loss: 0.0002315307647222653
step: 190, loss: 0.0005516273668035865
/usr/local/lib/python3.7/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=0.8304668304668306, f1=0.8337595907928389,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0072253113612532616
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.00018825425649993122
step: 20, loss: 0.0003494689881335944
step: 30, loss: 0.0014583944575861096
step: 40, loss: 0.001912241568788886
step: 50, loss: 0.002253051148727536

```

```

step: 60, loss: 0.001476187608204782
step: 70, loss: 0.0009532948606647551
step: 80, loss: 7.796590216457844e-05
step: 90, loss: 0.0013773003593087196
step: 100, loss: 0.00021450901112984866
step: 110, loss: 0.00304462225176394
step: 120, loss: 0.00014472039765678346
step: 130, loss: 0.0036153574474155903
step: 140, loss: 0.018948310986161232
step: 150, loss: 0.00026657592388801277
step: 160, loss: 0.0002128937339875847
step: 170, loss: 7.832504343241453e-05
step: 180, loss: 5.87660979363136e-05
step: 190, loss: 0.0255234744399786
/usr/local/lib/python3.7/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=0.8235294117647058, f1=0.8455284552845528,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026374959270469844
step: 10, loss: 0.00022875827562529594
step: 20, loss: 0.0003964862262364477
step: 30, loss: 0.00021668709814548492
step: 40, loss: 0.0002521632704883814
step: 50, loss: 0.003840785939246416
step: 60, loss: 0.00232509383931756
step: 70, loss: 0.00016281232819892466
step: 80, loss: 0.00011142776202177629
step: 90, loss: 6.492544343927875e-05
step: 100, loss: 0.0003178496262989938
step: 110, loss: 0.00011529131734278053
step: 120, loss: 0.0010709440102800727
step: 130, loss: 0.0003099295718129724
step: 140, loss: 0.008497812785208225
step: 150, loss: 0.0003062024188693613
step: 160, loss: 0.0001283810706809163
step: 170, loss: 0.0004020064661744982
step: 180, loss: 0.00023850180150475353
step: 190, loss: 0.0006263049435801804
/usr/local/lib/python3.7/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=0.8333333333333334, f1=0.8496042216358839,
best_f1=0.8435754189944134
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005821309750899673
step: 10, loss: 0.0004559689259622246
step: 20, loss: 0.0008135998505167663
step: 30, loss: 0.011919530108571053
step: 40, loss: 0.0016494556330144405
step: 50, loss: 8.044744026847184e-05
step: 60, loss: 7.13382542016916e-05
step: 70, loss: 0.00012505028280429542
step: 80, loss: 0.00039057235699146986
step: 90, loss: 0.003450939664617181
step: 100, loss: 0.0017709850799292326
step: 110, loss: 0.003291927045211196
step: 120, loss: 0.0013241184642538428
step: 130, loss: 0.0001482185034547001
step: 140, loss: 0.0037165724206715822
step: 150, loss: 7.465408998541534e-05
step: 160, loss: 0.13664965331554413
step: 170, loss: 0.0004993080510757864
step: 180, loss: 9.233065793523565e-05
step: 190, loss: 0.00015885428001638502
/usr/local/lib/python3.7/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=0.8346055979643766, f1=0.851851851851852,
best_f1=0.8435754189944134
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 [00:10, 189.66it/s]
load_f1 = 0.8695652173913043
real_f1 = 0.875
733it [00:00, 3284.74it/s]/usr/local/lib/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, 167.30it/s]

```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

```

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

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

```
step: 0, loss: 0.4927980899810791
```

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

```
step: 10, loss: 0.506746768951416
```

```
step: 20, loss: 0.2683795988559723
```

```
step: 30, loss: 0.4324023127555847
```

```
step: 40, loss: 0.5834211111068726
```

```
step: 50, loss: 0.3139124810695648
```

```
step: 60, loss: 0.5098637342453003
```

```

step: 70, loss: 0.2941877543926239
step: 80, loss: 0.23739448189735413
step: 90, loss: 0.24043980240821838
step: 100, loss: 0.15081319212913513
step: 110, loss: 0.39856258034706116
step: 120, loss: 0.31101495027542114
step: 130, loss: 0.31016021966934204
step: 140, loss: 0.3728073835372925
step: 150, loss: 0.32457566261291504
step: 160, loss: 0.4091988503932953
step: 170, loss: 0.29086926579475403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20689655172413793, f1=0.21552205471803462,
best_f1=0.21552205471803462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3098354637622833
step: 10, loss: 0.502065896987915
step: 20, loss: 0.31471091508865356
step: 30, loss: 0.30217939615249634
step: 40, loss: 0.06906390935182571
step: 50, loss: 0.38798823952674866
step: 60, loss: 0.2103300541639328
step: 70, loss: 0.4627819061279297
step: 80, loss: 0.2075711488723755
step: 90, loss: 0.250030517578125
step: 100, loss: 0.5040295720100403
step: 110, loss: 0.2607443034648895
step: 120, loss: 0.12326718121767044
step: 130, loss: 0.4006756544113159
step: 140, loss: 0.20099148154258728
step: 150, loss: 0.1876859813928604
step: 160, loss: 0.1403547078371048
step: 170, loss: 0.13615286350250244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7226107226107227, f1=0.7710843373493976,
best_f1=0.7710843373493976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2151722013950348
step: 10, loss: 0.07751200348138809
step: 20, loss: 0.08528495579957962
step: 30, loss: 0.04514985904097557
step: 40, loss: 0.12010633945465088
step: 50, loss: 0.3254382610321045
step: 60, loss: 0.15003792941570282
step: 70, loss: 0.11698716878890991
step: 80, loss: 0.22861778736114502
step: 90, loss: 0.21312935650348663
step: 100, loss: 0.034943122416734695
step: 110, loss: 0.08378417044878006
step: 120, loss: 0.12836842238903046
step: 130, loss: 0.28926199674606323
step: 140, loss: 0.13006530702114105
step: 150, loss: 0.098650261759758
step: 160, loss: 0.09669459611177444
step: 170, loss: 0.11167342960834503
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8100686498855835, f1=0.8258928571428572,
best_f1=0.8258928571428572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0765453353524208
step: 10, loss: 0.05433601513504982
step: 20, loss: 0.06906924396753311
step: 30, loss: 0.11065566539764404
step: 40, loss: 0.040510278195142746
step: 50, loss: 0.024550259113311768
step: 60, loss: 0.25847327709198
step: 70, loss: 0.008706070482730865
step: 80, loss: 0.33313602209091187
step: 90, loss: 0.18621058762073517
step: 100, loss: 0.31286197900772095
step: 110, loss: 0.2126898467540741
step: 120, loss: 0.2427627444267273

```

```

step: 130, loss: 0.058609992265701294
step: 140, loss: 0.07541526108980179
step: 150, loss: 0.14573635160923004
step: 160, loss: 0.00519581139087677
step: 170, loss: 0.11885125935077667
/usr/local/lib/python3.7/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=0.8019559902200489, f1=0.828235294117647,
best_f1=0.8258928571428572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030220042914152145
step: 10, loss: 0.04993230849504471
step: 20, loss: 0.14744240045547485
step: 30, loss: 0.005016608163714409
step: 40, loss: 0.06700095534324646
step: 50, loss: 0.016223788261413574
step: 60, loss: 0.05584670975804329
step: 70, loss: 0.016683947294950485
step: 80, loss: 0.0008620894514024258
step: 90, loss: 0.014981183223426342
step: 100, loss: 0.02048044465482235
step: 110, loss: 0.138109490275383
step: 120, loss: 0.07183363288640976
step: 130, loss: 0.002650853479281068
step: 140, loss: 0.1531028151512146
step: 150, loss: 0.06517212837934494
step: 160, loss: 0.11224613338708878
step: 170, loss: 0.007258345372974873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8431876606683805, f1=0.865, best_f1=0.865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15723751485347748
step: 10, loss: 0.006759663112461567
step: 20, loss: 0.08517920970916748
step: 30, loss: 0.009142708033323288
step: 40, loss: 0.0014894322957843542
step: 50, loss: 0.0026381092611700296

```

```

step: 60, loss: 0.038307517766952515
step: 70, loss: 0.02016925998032093
step: 80, loss: 0.025154974311590195
step: 90, loss: 0.07124143093824387
step: 100, loss: 0.01437375321984291
step: 110, loss: 0.044591426849365234
step: 120, loss: 0.01812789961695671
step: 130, loss: 0.016258087009191513
step: 140, loss: 0.05924302339553833
step: 150, loss: 0.012889509089291096
step: 160, loss: 0.07298707962036133
step: 170, loss: 0.09422429651021957
/usr/local/lib/python3.7/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=0.8346456692913385, f1=0.8678304239401496, best_f1=0.865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024134287610650063
step: 10, loss: 0.04112136363983154
step: 20, loss: 0.06649848818778992
step: 30, loss: 0.1459527164697647
step: 40, loss: 0.0006739392993040383
step: 50, loss: 0.0024040043354034424
step: 60, loss: 0.02265765704214573
step: 70, loss: 0.006506824400275946
step: 80, loss: 0.09132082015275955
step: 90, loss: 0.03471163660287857
step: 100, loss: 0.0007333848625421524
step: 110, loss: 0.0052372668869793415
step: 120, loss: 0.004505984950810671
step: 130, loss: 0.003068433376029134
step: 140, loss: 0.02546403557062149
step: 150, loss: 0.2513065040111542
step: 160, loss: 0.012910578399896622
step: 170, loss: 0.03928692638874054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8457446808510638, f1=0.863157894736842,
best_f1=0.863157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031876340508461
step: 10, loss: 0.014349891804158688
step: 20, loss: 0.015205735340714455
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0005048703169450164
step: 40, loss: 0.0016221872065216303
step: 50, loss: 0.0009536448633298278
step: 60, loss: 0.001513222698122263
step: 70, loss: 0.055499959737062454
step: 80, loss: 0.0015073855174705386
step: 90, loss: 0.07435086369514465
step: 100, loss: 0.001854547648690641
step: 110, loss: 0.039785340428352356
step: 120, loss: 0.16412346065044403
step: 130, loss: 0.015198457054793835
step: 140, loss: 0.07512449473142624
step: 150, loss: 0.007908460684120655
step: 160, loss: 0.007866433821618557
step: 170, loss: 0.2661641240119934
/usr/local/lib/python3.7/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=0.837092731829574, f1=0.8656716417910448,
best_f1=0.863157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031331513077020645
step: 10, loss: 0.0026842367369681597
step: 20, loss: 0.005249055102467537
step: 30, loss: 0.13650935888290405
step: 40, loss: 0.01827574335038662
step: 50, loss: 0.000511219201143831
step: 60, loss: 0.015105051919817924
step: 70, loss: 0.022621184587478638
step: 80, loss: 0.0010884527582675219
step: 90, loss: 0.0628746822476387
step: 100, loss: 0.0028970325365662575
step: 110, loss: 0.018901729956269264
step: 120, loss: 0.035477980971336365
step: 130, loss: 0.10677433758974075
step: 140, loss: 0.010059287771582603
step: 150, loss: 0.1481744945049286
step: 160, loss: 0.0012282804818823934
step: 170, loss: 0.012609321624040604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8614609571788413, f1=0.8480392156862744,
best_f1=0.8480392156862744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001732461852952838
step: 10, loss: 0.0007770452066324651
step: 20, loss: 0.007123639807105064
step: 30, loss: 0.052801914513111115
step: 40, loss: 0.00835336372256279
step: 50, loss: 0.004934818483889103
step: 60, loss: 0.11990700662136078
step: 70, loss: 0.0014050822937861085
step: 80, loss: 0.0025253077037632465
step: 90, loss: 0.0020732220727950335
step: 100, loss: 0.08864671736955643
step: 110, loss: 0.007185428403317928
step: 120, loss: 0.0004529857251327485
step: 130, loss: 0.005438453517854214
step: 140, loss: 0.0006887318450026214
step: 150, loss: 0.0007454646402038634
step: 160, loss: 0.000451472878921777
step: 170, loss: 0.041370779275894165
/usr/local/lib/python3.7/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=0.8396946564885497, f1=0.8457711442786069,
best_f1=0.8480392156862744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02477400377392769
step: 10, loss: 0.00026119500398635864
step: 20, loss: 0.0005259381723590195
step: 30, loss: 0.024911176413297653
step: 40, loss: 0.00019356599659658968
step: 50, loss: 0.003404487855732441
step: 60, loss: 0.006508774124085903
step: 70, loss: 0.014744502492249012
step: 80, loss: 0.010667501017451286
step: 90, loss: 0.0015929752262309194
step: 100, loss: 0.021667638793587685
step: 110, loss: 0.03843063488602638

```

```

step: 120, loss: 0.000342807819833979
step: 130, loss: 0.0037507587112486362
step: 140, loss: 0.0021164279896765947
step: 150, loss: 0.002210471546277404
step: 160, loss: 0.004888392053544521
step: 170, loss: 0.004244806244969368
/usr/local/lib/python3.7/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=0.8563968668407311, f1=0.8571428571428572,
best_f1=0.8480392156862744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002088752808049321
step: 10, loss: 0.002662366023287177
step: 20, loss: 0.00025564388488419354
step: 30, loss: 0.04060085862874985
step: 40, loss: 0.001857794588431716
step: 50, loss: 0.021384939551353455
step: 60, loss: 0.0013295153621584177
step: 70, loss: 0.00012589061225298792
step: 80, loss: 0.00011943963181693107
step: 90, loss: 0.0023518933448940516
step: 100, loss: 0.00024096027482300997
step: 110, loss: 0.0008559920242987573
step: 120, loss: 0.02842424437403679
step: 130, loss: 0.003981643822044134
step: 140, loss: 0.0008789505809545517
step: 150, loss: 0.00023917178623378277
step: 160, loss: 0.002178216353058815
step: 170, loss: 0.00022056377201806754
/usr/local/lib/python3.7/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=0.8439897698209718, f1=0.8542713567839195,
best_f1=0.8480392156862744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027993781259283423
step: 10, loss: 0.00010470645065652207
step: 20, loss: 0.0005281917983666062
step: 30, loss: 0.08223573863506317
step: 40, loss: 0.011291075497865677
step: 50, loss: 0.007273172494024038
step: 60, loss: 0.0018048674101009965
step: 70, loss: 0.16961129009723663

```

```

step: 80, loss: 0.0038053905591368675
step: 90, loss: 0.0005355463363230228
step: 100, loss: 0.011047613807022572
step: 110, loss: 0.0010836946312338114
step: 120, loss: 0.024123454466462135
step: 130, loss: 8.826397970551625e-05
step: 140, loss: 0.007824594154953957
step: 150, loss: 0.0008675375720486045
step: 160, loss: 9.459521243115887e-05
step: 170, loss: 0.00012105962377972901
/usr/local/lib/python3.7/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=0.8431876606683805, f1=0.8673469387755103,
best_f1=0.8480392156862744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010429563699290156
step: 10, loss: 0.00016762386076152325
step: 20, loss: 0.0019091746071353555
step: 30, loss: 0.00016114673053380102
step: 40, loss: 0.00010565127740846947
step: 50, loss: 0.0008915156358852983
step: 60, loss: 0.04592052847146988
step: 70, loss: 0.00022650834580417722
step: 80, loss: 0.00015852306387387216
step: 90, loss: 0.00015015651297289878
step: 100, loss: 0.009401382878422737
step: 110, loss: 0.017207255586981773
step: 120, loss: 8.41872242745012e-05
step: 130, loss: 0.020978428423404694
step: 140, loss: 0.01890026591718197
step: 150, loss: 0.01644166372716427
step: 160, loss: 0.0004512538725975901
step: 170, loss: 0.034649647772312164
/usr/local/lib/python3.7/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=0.8496042216358839, f1=0.8578811369509044,
best_f1=0.8480392156862744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.058020178694278e-05
step: 10, loss: 0.000382648897357285
step: 20, loss: 0.00010371991083957255
step: 30, loss: 0.0004079201607964933

```

```

step: 40, loss: 8.79575964063406e-05
step: 50, loss: 0.0002533577790018171
step: 60, loss: 0.00011132531653856859
step: 70, loss: 0.07320334017276764
step: 80, loss: 0.0001092962411348708
step: 90, loss: 0.002780991606414318
step: 100, loss: 0.014539874158799648
step: 110, loss: 0.0011272523552179337
step: 120, loss: 0.00022086936223786324
step: 130, loss: 0.00850137323141098
step: 140, loss: 0.003338359296321869
step: 150, loss: 8.536434324923903e-05
step: 160, loss: 0.00010452017158968374
step: 170, loss: 8.21919966256246e-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=0.8534031413612565, f1=0.8637532133676092,
best_f1=0.8480392156862744
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, 220.19it/s]
load_f1 = 0.8585858585858587
real_f1 = 0.8614609571788413
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 [00:33, 132.94it/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: 0.5871580243110657
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.5137525796890259
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: 0.4891977310180664
step: 30, loss: 0.2751196622848511
step: 40, loss: 0.35520192980766296
step: 50, loss: 0.5955165028572083
step: 60, loss: 0.46071359515190125
step: 70, loss: 0.24552886188030243
step: 80, loss: 0.25314459204673767
step: 90, loss: 0.27322879433631897
step: 100, loss: 0.08137717843055725
step: 110, loss: 0.15810725092887878
step: 120, loss: 0.10213150084018707
step: 130, loss: 0.007346839644014835
step: 140, loss: 0.039301976561546326
step: 150, loss: 0.18074433505535126
step: 160, loss: 0.034220825880765915
step: 170, loss: 0.45716696977615356
step: 180, loss: 0.07460378110408783
step: 190, loss: 0.12976565957069397
step: 200, loss: 0.0470656156539917
step: 210, loss: 0.030541911721229553
step: 220, loss: 0.06526480615139008
step: 230, loss: 0.006570532917976379

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9406494960806272, f1=0.9479638009049774,
best_f1=0.9479638009049774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008920794352889061
step: 10, loss: 0.08796384185552597
step: 20, loss: 0.0260833278298378
step: 30, loss: 0.02411065623164177
step: 40, loss: 0.11939213424921036
step: 50, loss: 0.012849947437644005
step: 60, loss: 0.20185662806034088
step: 70, loss: 0.0223771333694458
step: 80, loss: 0.0026332526467740536
step: 90, loss: 0.0440317802131176
step: 100, loss: 0.006831542123109102
step: 110, loss: 0.038581475615501404
step: 120, loss: 0.042563870549201965
step: 130, loss: 0.034924183040857315
step: 140, loss: 0.006286506541073322
step: 150, loss: 0.11001506447792053
step: 160, loss: 0.034149229526519775
step: 170, loss: 0.004080904182046652
step: 180, loss: 0.030607689172029495
step: 190, loss: 0.11431781202554703
step: 200, loss: 0.021384285762906075
step: 210, loss: 0.03007636032998562
step: 220, loss: 0.0022699208930134773
step: 230, loss: 0.0018830448389053345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9580973952434882, f1=0.9635535307517085,
best_f1=0.9635535307517085
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0070838541723787785
step: 10, loss: 0.009043838828802109
step: 20, loss: 0.026878954842686653
step: 30, loss: 0.01494133286178112
step: 40, loss: 0.0059510208666324615
step: 50, loss: 0.03934796527028084
step: 60, loss: 0.050901491194963455
step: 70, loss: 0.013216632418334484
step: 80, loss: 0.055474162101745605
step: 90, loss: 0.013955998234450817
step: 100, loss: 0.01365667860955
step: 110, loss: 0.02340252511203289
step: 120, loss: 0.0023278738372027874
step: 130, loss: 0.024031780660152435
step: 140, loss: 0.0017260591266676784
step: 150, loss: 0.2440161406993866
step: 160, loss: 0.011384967714548111
step: 170, loss: 0.0022052687127143145
step: 180, loss: 0.030777446925640106
step: 190, loss: 0.03370775282382965
step: 200, loss: 0.040587931871414185
step: 210, loss: 0.004919370170682669
step: 220, loss: 0.06871994584798813
step: 230, loss: 0.013214179314672947
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9594594594594594, f1=0.963963963963964,
best_f1=0.963963963963964
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014786751940846443
step: 10, loss: 0.005281464662402868
step: 20, loss: 0.004254079423844814
step: 30, loss: 0.008529792539775372
step: 40, loss: 0.12511146068572998
step: 50, loss: 0.024047890678048134
step: 60, loss: 0.009002722799777985
step: 70, loss: 0.02278275415301323
step: 80, loss: 0.10580209642648697
step: 90, loss: 0.04931539297103882
step: 100, loss: 0.020857729017734528
step: 110, loss: 0.001642333809286356

```



```

step: 120, loss: 0.006347838789224625
step: 130, loss: 0.0072830067947506905
step: 140, loss: 0.010602159425616264
step: 150, loss: 0.002740278374403715
step: 160, loss: 0.007970079779624939
step: 170, loss: 0.0035098684020340443
step: 180, loss: 0.0995093509554863
step: 190, loss: 0.004638928920030594
step: 200, loss: 0.047722455114126205
step: 210, loss: 0.07164303213357925
step: 220, loss: 0.0014948055613785982
step: 230, loss: 0.006435055751353502
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9715585893060296, f1=0.9679633867276889,
best_f1=0.9679633867276889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030482877045869827
step: 10, loss: 0.004116538446396589
step: 20, loss: 0.01380497869104147
step: 30, loss: 0.010803724639117718
step: 40, loss: 0.005996274296194315
step: 50, loss: 0.004333447199314833
step: 60, loss: 0.001763079664669931
step: 70, loss: 0.0020985575392842293
step: 80, loss: 0.22681088745594025
step: 90, loss: 0.10256848484277725
step: 100, loss: 0.0006381003186106682
step: 110, loss: 0.0024785767309367657
step: 120, loss: 0.0037858341820538044
step: 130, loss: 0.0005777380429208279
step: 140, loss: 0.004508154932409525
step: 150, loss: 0.011732647195458412
step: 160, loss: 0.00540114613249898
step: 170, loss: 0.0034251476172357798
step: 180, loss: 0.0034638661891222
step: 190, loss: 0.01849030703306198
step: 200, loss: 0.003394614439457655
step: 210, loss: 0.002361403312534094
step: 220, loss: 0.0019934538286179304
step: 230, loss: 0.017496280372142792

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9718785151856018, f1=0.9784335981838819,
best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009893302340060472
step: 10, loss: 0.0035584927536547184
step: 20, loss: 0.015392660163342953
step: 30, loss: 0.0008991702925413847
step: 40, loss: 0.00044784456258639693
step: 50, loss: 0.0013815519632771611
step: 60, loss: 0.005987759679555893
step: 70, loss: 0.0014661593595519662
step: 80, loss: 0.002065319800749421
step: 90, loss: 0.004806894343346357
step: 100, loss: 0.0021124309860169888
step: 110, loss: 0.005960024893283844
step: 120, loss: 0.0006490078521892428
step: 130, loss: 0.0010664473520591855
step: 140, loss: 0.033939577639102936
step: 150, loss: 0.00042402197141200304
step: 160, loss: 0.003231536829844117
step: 170, loss: 0.0013806025963276625
step: 180, loss: 0.001987106166779995
step: 190, loss: 0.00282041123136878
step: 200, loss: 0.0028496619779616594
step: 210, loss: 0.0007488343981094658
step: 220, loss: 0.09240757673978806
step: 230, loss: 0.00474504753947258
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9807474518686297, f1=0.9750566893424036,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00398636981844902
step: 10, loss: 0.0019331941148266196
step: 20, loss: 0.0007581542595289648
step: 30, loss: 0.0004980515805073082
step: 40, loss: 0.001130406279116869
step: 50, loss: 0.0012665078975260258
step: 60, loss: 0.001318678492680192
step: 70, loss: 0.0005855453200638294
step: 80, loss: 0.0009164768853224814
step: 90, loss: 0.00041508369031362236
step: 100, loss: 0.0003521550679579377
step: 110, loss: 0.0005000985111109912
step: 120, loss: 0.0006148521788418293
step: 130, loss: 0.12910473346710205
step: 140, loss: 0.0006902384920977056
step: 150, loss: 0.21706551313400269
step: 160, loss: 0.021126531064510345
step: 170, loss: 0.00489415368065238
step: 180, loss: 0.0023095114156603813
step: 190, loss: 0.018550600856542587
step: 200, loss: 0.0007655165391042829
step: 210, loss: 0.06597058475017548
step: 220, loss: 0.0007042507641017437
step: 230, loss: 0.001331223757006228
/usr/local/lib/python3.7/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=0.9766407119021134, f1=0.967525195968645,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010274696396663785
step: 10, loss: 0.0034720548428595066
step: 20, loss: 0.003302976954728365
step: 30, loss: 0.0013923302758485079
step: 40, loss: 0.002282612957060337
step: 50, loss: 0.0010143170366063714
step: 60, loss: 0.0015508889919146895
step: 70, loss: 0.00041776237776502967
step: 80, loss: 0.022057926282286644
step: 90, loss: 0.0008892167243175209
step: 100, loss: 0.0012389144394546747
step: 110, loss: 0.0006190964486449957
step: 120, loss: 0.0005163397872820497
step: 130, loss: 0.0009948507649824023
step: 140, loss: 0.0005248537636362016
step: 150, loss: 0.16841110587120056

```

```

step: 160, loss: 0.000882919819559902
step: 170, loss: 0.005936041008681059
step: 180, loss: 0.0005744318477809429
step: 190, loss: 0.002584804780781269
step: 200, loss: 0.006113177631050348
step: 210, loss: 0.016543185338377953
step: 220, loss: 0.003713525366038084
step: 230, loss: 0.0007449750555679202
/usr/local/lib/python3.7/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=0.9787234042553192, f1=0.976324689966178,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012046161573380232
step: 10, loss: 0.07617294043302536
step: 20, loss: 0.0017328757094219327
step: 30, loss: 0.002088341396301985
step: 40, loss: 0.001452173339203
step: 50, loss: 0.0012985066277906299
step: 60, loss: 0.0014718278544023633
step: 70, loss: 0.015908606350421906
step: 80, loss: 0.0006997567252255976
step: 90, loss: 0.024734269827604294
step: 100, loss: 0.0010178226511925459
step: 110, loss: 0.00021951855160295963
step: 120, loss: 0.0008242623880505562
step: 130, loss: 0.0010992089519277215
step: 140, loss: 0.001246394356712699
step: 150, loss: 0.07859258353710175
step: 160, loss: 0.012821673415601254
step: 170, loss: 0.0006191351567395031
step: 180, loss: 0.02071942202746868
step: 190, loss: 0.00039985880721360445
step: 200, loss: 0.0011494260979816318
step: 210, loss: 0.036492686718702316
step: 220, loss: 0.0006233828607946634
step: 230, loss: 0.0004282299196347594
/usr/local/lib/python3.7/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=0.9784335981838819, f1=0.9739524348810873,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0006717591895721853
step: 10, loss: 0.0015180568443611264
step: 20, loss: 0.0009404206648468971
step: 30, loss: 0.00028674150235019624
step: 40, loss: 0.0012667254777625203
step: 50, loss: 0.00028033062699250877
step: 60, loss: 0.00051696109585464
step: 70, loss: 0.0048556639812886715
step: 80, loss: 0.01605396158993244
step: 90, loss: 0.0006221953080967069
step: 100, loss: 0.0012267997954040766
step: 110, loss: 0.004075950477272272
step: 120, loss: 0.00041695949039421976
step: 130, loss: 0.00225185533054173
step: 140, loss: 0.0005793216405436397
step: 150, loss: 0.0016701502026990056
step: 160, loss: 0.00027928908821195364
step: 170, loss: 0.0004114377370569855
step: 180, loss: 0.0011615591356530786
step: 190, loss: 0.0006311158649623394
step: 200, loss: 0.012393683195114136
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 210, loss: 0.0014624539762735367
step: 220, loss: 0.0008750060806050897
step: 230, loss: 0.000366578227840364
/usr/local/lib/python3.7/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=0.9773242630385486, f1=0.9751131221719457,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046319447574205697
step: 10, loss: 0.002425061073154211
step: 20, loss: 0.0004389541572891176
step: 30, loss: 0.0008519343100488186
step: 40, loss: 0.00011687118239933625
step: 50, loss: 0.00019929588597733527
step: 60, loss: 0.018143180757761
step: 70, loss: 0.0004620740655809641
step: 80, loss: 0.006663662847131491
step: 90, loss: 0.20631712675094604
step: 100, loss: 0.00130231655202806
step: 110, loss: 0.0011630502995103598
step: 120, loss: 0.0009255344630219042
step: 130, loss: 0.0006694981129840016
step: 140, loss: 0.009498958475887775

```

```

step: 150, loss: 0.003711393568664789
step: 160, loss: 0.008531942032277584
step: 170, loss: 0.0009946717182174325
step: 180, loss: 0.0008585212053731084
step: 190, loss: 0.0005529208574444056
step: 200, loss: 0.0012050268705934286
step: 210, loss: 0.0007207072922028601
step: 220, loss: 0.0011683411430567503
step: 230, loss: 0.0008193366229534149
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.980963045912654, f1=0.9684684684684683,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004547573626041412
step: 10, loss: 0.0036095979157835245
step: 20, loss: 0.004503483418375254
step: 30, loss: 0.13923247158527374
step: 40, loss: 0.0010060588829219341
step: 50, loss: 0.002389013534411788
step: 60, loss: 0.0011616303818300366
step: 70, loss: 0.0009989994578063488
step: 80, loss: 0.00019388018699828535
step: 90, loss: 0.0008884332492016256
step: 100, loss: 0.00023743242491036654
step: 110, loss: 0.00021577210281975567
step: 120, loss: 0.0004494485619943589
step: 130, loss: 0.00039662656490691006
step: 140, loss: 0.00094194570556283
step: 150, loss: 0.0007643340504728258
step: 160, loss: 0.002842253539711237
step: 170, loss: 0.0007476579048670828
step: 180, loss: 0.0007625713478773832
step: 190, loss: 0.001397319370880723
step: 200, loss: 0.0006303559057414532
step: 210, loss: 0.0014670079108327627
step: 220, loss: 0.0003020043659489602
step: 230, loss: 0.0009369557956233621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841628959276018, f1=0.971815107102593,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015509293880313635
step: 10, loss: 0.0006351436022669077
step: 20, loss: 0.0015771619509905577
step: 30, loss: 0.00030298996716737747
step: 40, loss: 0.0017216593259945512
step: 50, loss: 0.004516018554568291
step: 60, loss: 0.0006756550865247846
step: 70, loss: 0.0005544899613596499
step: 80, loss: 0.00037972588324919343
step: 90, loss: 0.0007459760527126491
step: 100, loss: 0.001282707555219531
step: 110, loss: 0.002046257024630904
step: 120, loss: 0.0010034984443336725
step: 130, loss: 0.0009079648298211396
step: 140, loss: 0.00027704695821739733
step: 150, loss: 0.0002696705050766468
step: 160, loss: 0.0016784326871857047
step: 170, loss: 0.0014357691397890449
step: 180, loss: 0.06384280323982239
step: 190, loss: 0.0011806460097432137
step: 200, loss: 0.00011258203448960558
step: 210, loss: 0.004909297451376915
step: 220, loss: 0.0006586180534213781
step: 230, loss: 0.0013112318702042103
/usr/local/lib/python3.7/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=0.9771689497716894, f1=0.9715585893060296,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004732392553705722
step: 10, loss: 0.0003285110287833959
step: 20, loss: 0.0004778598085977137
step: 30, loss: 0.0007353800465352833
step: 40, loss: 0.0007251208880916238
step: 50, loss: 0.00028160272631794214
step: 60, loss: 0.0006174191366881132

```

```

step: 70, loss: 0.000519675319083035
step: 80, loss: 0.00035361392656341195
step: 90, loss: 0.0012431818759068847
step: 100, loss: 0.0013918336480855942
step: 110, loss: 0.007636246737092733
step: 120, loss: 0.00011211410310352221
step: 130, loss: 0.0016210338799282908
step: 140, loss: 0.0006332962075248361
step: 150, loss: 0.0002940078265964985
step: 160, loss: 0.00194832356646657
step: 170, loss: 0.0027006168384104967
step: 180, loss: 0.0005213369731791317
step: 190, loss: 0.0014479877427220345
step: 200, loss: 0.01675317995250225
step: 210, loss: 0.0005290696863085032
step: 220, loss: 0.0010552993044257164
step: 230, loss: 0.001694210572168231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, 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: 0.006960004102438688
step: 10, loss: 0.0005601196899078786
step: 20, loss: 0.001298147370107472
step: 30, loss: 0.0008208613726310432
step: 40, loss: 0.0004048601258546114
step: 50, loss: 0.000653019524179399
step: 60, loss: 0.03142153471708298
step: 70, loss: 0.00034919779864139855
step: 80, loss: 0.0007173779886215925
step: 90, loss: 0.0003009735664818436
step: 100, loss: 0.0002733810106292367
step: 110, loss: 0.00046177173499017954
step: 120, loss: 0.044298261404037476
step: 130, loss: 0.0009306760039180517
step: 140, loss: 0.010914687998592854
step: 150, loss: 0.0007622306584380567
step: 160, loss: 0.021417811512947083
step: 170, loss: 0.00024064292665570974
step: 180, loss: 0.0005873176851309836

```



```

step: 190, loss: 0.0005602984456345439
step: 200, loss: 0.0021216229069978
step: 210, loss: 0.02283608168363571
step: 220, loss: 0.0006829475751146674
step: 230, loss: 0.0003482132451608777
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9696969696969697,
best_f1=0.9698324022346367
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:17, 143.03it/s]
load_f1 = 0.9830890642615557
real_f1 = 0.9831649831649831
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, 134.22it/s]

```

2.3.2 DBLP-GoogleScholar - 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-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/
```

```
Downloading: 100% 481/481 [00:00<00:00, 497kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 2.40MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 1.79MB/s]
Downloading: 100% 501M/501M [00:16<00:00, 30.2MB/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               : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale              : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6532546281814575
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: 0.47824540734291077
step: 20, loss: 0.2524379789829254
step: 30, loss: 0.3680313229560852
step: 40, loss: 0.3169473707675934
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.15404440462589264
step: 60, loss: 0.18225400149822235
step: 70, loss: 0.21142041683197021
step: 80, loss: 0.0923062190413475
step: 90, loss: 0.17734487354755402
step: 100, loss: 0.4316587746143341
step: 110, loss: 0.10905889421701431
step: 120, loss: 0.10502088069915771
step: 130, loss: 0.1339646577835083
step: 140, loss: 0.2741522490978241
step: 150, loss: 0.12413064390420914
step: 160, loss: 0.15160778164863586
step: 170, loss: 0.08497345447540283
step: 180, loss: 0.06989573687314987
step: 190, loss: 0.08201098442077637
step: 200, loss: 0.07049285620450974
step: 210, loss: 0.05291526019573212
step: 220, loss: 0.02953203022480011
step: 230, loss: 0.19211028516292572
step: 240, loss: 0.06543515622615814
step: 250, loss: 0.09677764773368835
step: 260, loss: 0.1388375163078308
step: 270, loss: 0.6789646744728088
step: 280, loss: 0.08033513277769089
step: 290, loss: 0.20176155865192413
step: 300, loss: 0.02639298513531685
step: 310, loss: 0.09051183611154556

```

```

step: 320, loss: 0.014086684212088585
step: 330, loss: 0.0860525444149971
step: 340, loss: 0.293308287858963
step: 350, loss: 0.056265994906425476
step: 360, loss: 0.025184454396367073
step: 370, loss: 0.03325507044792175
step: 380, loss: 0.13695374131202698
step: 390, loss: 0.03258060663938522
step: 400, loss: 0.029831692576408386
step: 410, loss: 0.20999041199684143
step: 420, loss: 0.07050420343875885
step: 430, loss: 0.026992250233888626
step: 440, loss: 0.12033317238092422
step: 450, loss: 0.02119775116443634
step: 460, loss: 0.05620047077536583
step: 470, loss: 0.042656175792217255
step: 480, loss: 0.1879993975162506
step: 490, loss: 0.07329455763101578
step: 500, loss: 0.01000463031232357
step: 510, loss: 0.0817284807562828
step: 520, loss: 0.05483602359890938
step: 530, loss: 0.005617056041955948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9404096834264433, f1=0.9356943150046597,
best_f1=0.9356943150046597
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06734523922204971
step: 10, loss: 0.06944850087165833
step: 20, loss: 0.01362349558621645
step: 30, loss: 0.07850877195596695
step: 40, loss: 0.03618954122066498
step: 50, loss: 0.2100202590227127
step: 60, loss: 0.017648985609412193
step: 70, loss: 0.022478733211755753
step: 80, loss: 0.014390720054507256
step: 90, loss: 0.03749530762434006
step: 100, loss: 0.17699532210826874
step: 110, loss: 0.020124349743127823
step: 120, loss: 0.02086193487048149
step: 130, loss: 0.010215030051767826

```

```

step: 140, loss: 0.009296794421970844
step: 150, loss: 0.006571830250322819
step: 160, loss: 0.05508023872971535
step: 170, loss: 0.062434785068035126
step: 180, loss: 0.013390779495239258
step: 190, loss: 0.010442554950714111
step: 200, loss: 0.09740457683801651
step: 210, loss: 0.046637482941150665
step: 220, loss: 0.001732812961563468
step: 230, loss: 0.049055177718400955
step: 240, loss: 0.004089142195880413
step: 250, loss: 0.09795358777046204
step: 260, loss: 0.08662408590316772
step: 270, loss: 0.01583450473845005
step: 280, loss: 0.054341014474630356
step: 290, loss: 0.03470909968018532
step: 300, loss: 0.032874707132577896
step: 310, loss: 0.033955249935388565
step: 320, loss: 0.05089721828699112
step: 330, loss: 0.08009182661771774
step: 340, loss: 0.12209313362836838
step: 350, loss: 0.0018624719232320786
step: 360, loss: 0.19213466346263885
step: 370, loss: 0.01780455932021141
step: 380, loss: 0.06833265721797943
step: 390, loss: 0.003815979929640889
step: 400, loss: 0.07900500297546387
step: 410, loss: 0.09932415932416916
step: 420, loss: 0.05981418490409851
step: 430, loss: 0.17841637134552002
step: 440, loss: 0.04170027747750282
step: 450, loss: 0.031586553901433945
step: 460, loss: 0.0232792180031538
step: 470, loss: 0.09915716201066971
step: 480, loss: 0.1396629810333252
step: 490, loss: 0.03206980600953102
step: 500, loss: 0.009015994146466255
step: 510, loss: 0.008247480727732182
step: 520, loss: 0.29032430052757263
step: 530, loss: 0.1676531732082367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9480093676814988, f1=0.9449112978524743,

```

```
best_f1=0.9449112978524743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0883052870631218
step: 10, loss: 0.04604554548859596
step: 20, loss: 0.0058432212099432945
step: 30, loss: 0.027056924998760223
step: 40, loss: 0.05964023619890213
step: 50, loss: 0.011109530925750732
step: 60, loss: 0.010546381585299969
step: 70, loss: 0.031130356714129448
step: 80, loss: 0.023050358518958092
step: 90, loss: 0.15482570230960846
step: 100, loss: 0.05834965780377388
step: 110, loss: 0.09169831871986389
step: 120, loss: 0.09687283635139465
step: 130, loss: 0.04602190479636192
step: 140, loss: 0.0051700640469789505
step: 150, loss: 0.02024921216070652
step: 160, loss: 0.025809505954384804
step: 170, loss: 0.008748043328523636
step: 180, loss: 0.011726034805178642
step: 190, loss: 0.0044944072142243385
step: 200, loss: 0.08966061472892761
step: 210, loss: 0.02815387025475502
step: 220, loss: 0.06527848541736603
step: 230, loss: 0.10657674074172974
step: 240, loss: 0.07207324355840683
step: 250, loss: 0.03586428239941597
step: 260, loss: 0.02261986769735813
step: 270, loss: 0.005980911664664745
step: 280, loss: 0.00023538109962828457
step: 290, loss: 0.007817242294549942
step: 300, loss: 0.088372603058815
step: 310, loss: 0.12347527593374252
step: 320, loss: 0.10007532685995102
step: 330, loss: 0.04566723480820656
step: 340, loss: 0.0043404302559792995
step: 350, loss: 0.042401861399412155
step: 360, loss: 0.011329410597682
step: 370, loss: 0.026797812432050705
step: 380, loss: 0.13879474997520447
step: 390, loss: 0.011184110306203365
step: 400, loss: 0.11748795956373215
step: 410, loss: 0.03660549968481064
step: 420, loss: 0.003108459524810314
step: 430, loss: 0.03153286129236221
```

```

step: 440, loss: 0.20164120197296143
step: 450, loss: 0.05019167810678482
step: 460, loss: 0.14581626653671265
step: 470, loss: 0.07109794020652771
step: 480, loss: 0.2241692692041397
step: 490, loss: 0.06597208231687546
step: 500, loss: 0.07282087206840515
step: 510, loss: 0.006146302446722984
step: 520, loss: 0.10133993625640869
step: 530, loss: 0.011293089017271996
/usr/local/lib/python3.7/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=0.9449626865671642, f1=0.927468413664015,
best_f1=0.9449112978524743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02576279453933239
step: 10, loss: 0.0015649001579731703
step: 20, loss: 0.10998547077178955
step: 30, loss: 0.013971655629575253
step: 40, loss: 0.005784517154097557
step: 50, loss: 0.009484986774623394
step: 60, loss: 0.020684074610471725
step: 70, loss: 0.04722961038351059
step: 80, loss: 0.2706667482852936
step: 90, loss: 0.007970619946718216
step: 100, loss: 0.011373440735042095
step: 110, loss: 0.11805833131074905
step: 120, loss: 0.002184451324865222
step: 130, loss: 0.011055530980229378
step: 140, loss: 0.016601532697677612
step: 150, loss: 0.009311995469033718
step: 160, loss: 0.03170617297291756
step: 170, loss: 0.007080815266817808
step: 180, loss: 0.01963677816092968
step: 190, loss: 0.021039247512817383
step: 200, loss: 0.1182284876704216
step: 210, loss: 0.0013173534534871578
step: 220, loss: 0.004711833782494068
step: 230, loss: 0.003787215333431959
step: 240, loss: 0.029755622148513794
step: 250, loss: 0.03570142015814781
step: 260, loss: 0.0008566994220018387
step: 270, loss: 0.05531091243028641
step: 280, loss: 0.0011357659241184592
step: 290, loss: 0.011754270642995834

```

```

step: 300, loss: 0.013950970023870468
step: 310, loss: 0.0054148719646036625
step: 320, loss: 0.04207371175289154
step: 330, loss: 0.049757130444049835
step: 340, loss: 0.012554898858070374
step: 350, loss: 0.05497339367866516
step: 360, loss: 0.03812260180711746
step: 370, loss: 0.00836033746600151
step: 380, loss: 0.007937065325677395
step: 390, loss: 0.001970376819372177
step: 400, loss: 0.09692850708961487
step: 410, loss: 0.0012660013744607568
step: 420, loss: 0.026671336963772774
step: 430, loss: 0.0037119623739272356
step: 440, loss: 0.0019150603329762816
step: 450, loss: 0.014145140536129475
step: 460, loss: 0.012215581722557545
step: 470, loss: 0.002825212897732854
step: 480, loss: 0.0019384390907362103
step: 490, loss: 0.0018622392090037465
step: 500, loss: 0.07079384475946426
step: 510, loss: 0.10192796587944031
step: 520, loss: 0.03667086735367775
step: 530, loss: 0.10950618982315063
/usr/local/lib/python3.7/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=0.9478098788443615, f1=0.9364269141531322,
best_f1=0.9449112978524743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002934279851615429
step: 10, loss: 0.024892445653676987
step: 20, loss: 0.0080583356320858
step: 30, loss: 0.002573499921709299
step: 40, loss: 0.02644769288599491
step: 50, loss: 0.004040106665343046
step: 60, loss: 0.018426625058054924
step: 70, loss: 0.030998343601822853
step: 80, loss: 0.00030314933974295855
step: 90, loss: 0.026889581233263016
step: 100, loss: 0.02116321213543415
step: 110, loss: 0.0006783325807191432
step: 120, loss: 0.043043553829193115
step: 130, loss: 0.004043241497129202
step: 140, loss: 0.0020993186626583338
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

```



```

step: 150, loss: 0.018275970593094826
step: 160, loss: 0.0015142710180953145
step: 170, loss: 0.044713180512189865
step: 180, loss: 0.005108017940074205
step: 190, loss: 0.0059374249540269375
step: 200, loss: 0.004166995640844107
step: 210, loss: 0.000625689048320055
step: 220, loss: 0.005350057035684586
step: 230, loss: 0.0017139199189841747
step: 240, loss: 0.02534278854727745
step: 250, loss: 0.17447295784950256
step: 260, loss: 0.0011129140621051192
step: 270, loss: 0.017017552629113197
step: 280, loss: 0.00623171916231513
step: 290, loss: 0.0009003037703223526
step: 300, loss: 0.10096976161003113
step: 310, loss: 0.06043640524148941
step: 320, loss: 0.05122688040137291
step: 330, loss: 0.004172894172370434
step: 340, loss: 0.0008124663727357984
step: 350, loss: 0.0007996626081876457
step: 360, loss: 0.0001376587460981682
step: 370, loss: 0.000287405593553558
step: 380, loss: 0.0001258361153304577
step: 390, loss: 0.0013938165502622724
step: 400, loss: 0.11884628981351852
step: 410, loss: 0.08342193067073822
step: 420, loss: 0.10634318739175797
step: 430, loss: 0.003078061854466796
step: 440, loss: 0.0032855828758329153
step: 450, loss: 0.03467727452516556
step: 460, loss: 0.01564655266702175
step: 470, loss: 0.04946082457900047
step: 480, loss: 0.008940781466662884
step: 490, loss: 0.01535582635551691
step: 500, loss: 0.09345578402280807
step: 510, loss: 0.0007502530352212489
step: 520, loss: 0.10918381810188293
step: 530, loss: 0.02906683459877968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9491525423728814, f1=0.9383280036546369,
best_f1=0.9383280036546369

```

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

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

step: 0, loss: 0.00979416910558939
step: 10, loss: 0.0007662491989322007
step: 20, loss: 0.016550308093428612
step: 30, loss: 0.0009387903846800327
step: 40, loss: 0.001012665219604969
step: 50, loss: 0.00029806740349158645
step: 60, loss: 0.0010948039125651121
step: 70, loss: 0.00017072477203328162
step: 80, loss: 0.00179264135658741
step: 90, loss: 0.05069095641374588
step: 100, loss: 0.007713434286415577
step: 110, loss: 0.005837816745042801
step: 120, loss: 0.01054073590785265
step: 130, loss: 0.001562771969474852
step: 140, loss: 0.0013491937424987555
step: 150, loss: 0.05149867385625839
step: 160, loss: 0.026731085032224655
step: 170, loss: 0.008410760201513767
step: 180, loss: 0.0018638247856870294
step: 190, loss: 0.23485396802425385
step: 200, loss: 0.00251529342494905
step: 210, loss: 0.01966635137796402
step: 220, loss: 0.003461054991930723
step: 230, loss: 0.014838453382253647
step: 240, loss: 0.008815902285277843
step: 250, loss: 0.053515657782554626
step: 260, loss: 0.0015350221656262875
step: 270, loss: 0.008494212292134762
step: 280, loss: 0.0029202818404883146
step: 290, loss: 0.0004882964421994984
step: 300, loss: 0.06626340746879578
step: 310, loss: 0.03641403093934059
step: 320, loss: 0.0002054188516922295
step: 330, loss: 0.004445885308086872
step: 340, loss: 0.00015817908570170403
step: 350, loss: 0.0035279039293527603
step: 360, loss: 0.18121758103370667
step: 370, loss: 0.01610967516899109
step: 380, loss: 0.0013016827870160341
step: 390, loss: 0.0030486066825687885
step: 400, loss: 0.039905909448862076
step: 410, loss: 0.006648289039731026
step: 420, loss: 0.001974803628399968
step: 430, loss: 0.0008344577508978546
step: 440, loss: 0.0006745218997821212

```
step: 450, loss: 0.01447316911071539
step: 460, loss: 0.0022127265110611916
step: 470, loss: 0.001261041616089642
step: 480, loss: 0.005463081412017345
step: 490, loss: 0.050794005393981934
step: 500, loss: 0.0529417023062706
step: 510, loss: 0.13139697909355164
step: 520, loss: 0.03135494142770767
step: 530, loss: 0.0009134431602433324
/usr/local/lib/python3.7/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=0.9460227272727273, f1=0.936470588235294,
best_f1=0.9383280036546369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02484842762351036
step: 10, loss: 0.0016410844400525093
step: 20, loss: 0.013581291772425175
step: 30, loss: 0.007544277235865593
step: 40, loss: 0.002971825422719121
step: 50, loss: 0.012031758204102516
step: 60, loss: 0.010240491479635239
step: 70, loss: 0.003555386560037732
step: 80, loss: 0.001005023019388318
step: 90, loss: 0.00021950469817966223
step: 100, loss: 0.0011805961839854717
step: 110, loss: 0.000429431238444522
step: 120, loss: 0.008810224942862988
step: 130, loss: 0.00039104610914364457
step: 140, loss: 0.001651870901696384
step: 150, loss: 0.006835335399955511
step: 160, loss: 0.00020102763664908707
step: 170, loss: 0.009639469906687737
step: 180, loss: 0.15726731717586517
step: 190, loss: 0.043224431574344635
step: 200, loss: 0.0005773542216047645
step: 210, loss: 0.0006361650303006172
step: 220, loss: 0.00020692500402219594
step: 230, loss: 0.0001417861203663051
step: 240, loss: 0.005632961168885231
step: 250, loss: 0.0010664894944056869
step: 260, loss: 0.02173565700650215
step: 270, loss: 0.00023436194169335067
step: 280, loss: 0.021015798673033714
step: 290, loss: 0.00021953447139821947
step: 300, loss: 0.00016758286801632494
```

```
step: 310, loss: 0.00011741340858861804
step: 320, loss: 0.007384579628705978
step: 330, loss: 5.4784686653874815e-05
step: 340, loss: 0.022548343986272812
step: 350, loss: 0.0016656587831676006
step: 360, loss: 0.0006491955136880279
step: 370, loss: 0.0021966390777379274
step: 380, loss: 0.015134756453335285
step: 390, loss: 0.00840156339108944
step: 400, loss: 0.07320155203342438
step: 410, loss: 0.02411503717303276
step: 420, loss: 0.006741268560290337
step: 430, loss: 0.0004737121344078332
step: 440, loss: 0.0011635632254183292
step: 450, loss: 0.08029985427856445
step: 460, loss: 0.0466441810131073
step: 470, loss: 0.18141774833202362
step: 480, loss: 0.004636579193174839
step: 490, loss: 0.0031595099717378616
step: 500, loss: 0.00106896017678082
step: 510, loss: 0.0003491169190965593
step: 520, loss: 0.004766672849655151
step: 530, loss: 0.004069885704666376
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9476609541454377, f1=0.9405255878284925,
best_f1=0.9383280036546369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000767586927395314
step: 10, loss: 0.0014422731474041939
step: 20, loss: 0.00022568779240828007
step: 30, loss: 0.03943060338497162
step: 40, loss: 4.731707667815499e-05
step: 50, loss: 0.001296322443522513
step: 60, loss: 0.00010718701378209516
step: 70, loss: 0.000308704242343083
step: 80, loss: 0.0002886533038690686
step: 90, loss: 0.0004181099066045135
step: 100, loss: 0.0003143748617731035
step: 110, loss: 5.376310582505539e-05
step: 120, loss: 5.0178419769508764e-05
step: 130, loss: 0.00043709605233743787
step: 140, loss: 7.418434688588604e-05
step: 150, loss: 5.562582373386249e-05
step: 160, loss: 2.8068678147974424e-05
```

```

step: 170, loss: 0.0039024604484438896
step: 180, loss: 0.0015010031638666987
step: 190, loss: 0.0021858015097677708
step: 200, loss: 0.009349007159471512
step: 210, loss: 0.11579433083534241
step: 220, loss: 0.000857498322147876
step: 230, loss: 0.07458625733852386
step: 240, loss: 0.00986485742032528
step: 250, loss: 0.000977838528342545
step: 260, loss: 0.0004401802143547684
step: 270, loss: 0.0017235854174941778
step: 280, loss: 0.0008204776095226407
step: 290, loss: 0.00028716641827486455
step: 300, loss: 0.00017344419029541314
step: 310, loss: 0.0008413958130404353
step: 320, loss: 0.009455080144107342
step: 330, loss: 0.000175258974195458
step: 340, loss: 0.01905435509979725
step: 350, loss: 0.0010461027268320322
step: 360, loss: 0.10364925861358643
step: 370, loss: 0.1861714869737625
step: 380, loss: 0.0005848844884894788
step: 390, loss: 0.002244841307401657
step: 400, loss: 0.00039949669735506177
step: 410, loss: 0.0017902359832078218
step: 420, loss: 0.00010757723794085905
step: 430, loss: 7.976838969625533e-05
step: 440, loss: 0.00013792562822345644
step: 450, loss: 8.943086140789092e-05
step: 460, loss: 0.0004424174258019775
step: 470, loss: 0.02708972431719303
step: 480, loss: 0.010070468299090862
step: 490, loss: 0.08720697462558746
step: 500, loss: 0.00169320625718683
step: 510, loss: 0.005030814558267593
step: 520, loss: 0.0008133291848935187
step: 530, loss: 0.0016135042533278465
/usr/local/lib/python3.7/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=0.9491211840888066, f1=0.9398806792106471,
best_f1=0.9383280036546369
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016376236453652382
step: 10, loss: 0.0006459199939854443
step: 20, loss: 0.0005534177762456238

```

step: 30, loss: 0.04792246222496033
step: 40, loss: 0.00016842079639900476
step: 50, loss: 0.00035528390435501933
step: 60, loss: 0.0009704524418339133
step: 70, loss: 0.0002797002671286464
step: 80, loss: 0.0016353853279724717
step: 90, loss: 0.05666283518075943
step: 100, loss: 0.0007751267403364182
step: 110, loss: 0.004409410525113344
step: 120, loss: 0.0001195806689793244
step: 130, loss: 0.0011967496247962117
step: 140, loss: 0.05886802449822426
step: 150, loss: 0.00048146297922357917
step: 160, loss: 4.3602725781966e-05
step: 170, loss: 0.0003179931081831455
step: 180, loss: 0.019961126148700714
step: 190, loss: 0.0001586825674166903
step: 200, loss: 0.0018723881803452969
step: 210, loss: 0.0005263371858745813
step: 220, loss: 4.803571937372908e-05
step: 230, loss: 7.236217061290517e-05
step: 240, loss: 7.317541167140007e-05
step: 250, loss: 3.206156543456018e-05
step: 260, loss: 5.698244058294222e-05
step: 270, loss: 0.0005706415395252407
step: 280, loss: 0.0001897416659630835
step: 290, loss: 2.3416088879457675e-05
step: 300, loss: 0.08377489447593689
step: 310, loss: 0.05314505100250244
step: 320, loss: 0.0004595532373059541
step: 330, loss: 0.000978741911239922
step: 340, loss: 0.020320050418376923
step: 350, loss: 0.014773596078157425
step: 360, loss: 0.04884704574942589
step: 370, loss: 0.0006795090739615262
step: 380, loss: 0.0014035619096830487
step: 390, loss: 0.0010920347413048148
step: 400, loss: 0.031159337610006332
step: 410, loss: 0.03025370091199875
step: 420, loss: 0.01773151010274887
step: 430, loss: 0.013810890726745129
step: 440, loss: 4.025211455882527e-05
step: 450, loss: 0.0020123689901083708
step: 460, loss: 0.00013361798482947052
step: 470, loss: 0.017741119489073753
step: 480, loss: 4.9615526222623885e-05
step: 490, loss: 0.0009374269284307957
step: 500, loss: 5.183240500628017e-05

```

step: 510, loss: 0.022974418476223946
step: 520, loss: 0.002424411941319704
step: 530, loss: 0.019108261913061142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9541627689429373, f1=0.9424326833797586,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003758902312256396
step: 10, loss: 4.718622221844271e-05
step: 20, loss: 0.0016835584538057446
step: 30, loss: 0.00022987548436503857
step: 40, loss: 5.17990774824284e-05
step: 50, loss: 0.0001019513001665473
step: 60, loss: 6.72384921927005e-05
step: 70, loss: 2.4194701836677268e-05
step: 80, loss: 0.004498879425227642
step: 90, loss: 5.260364559944719e-05
step: 100, loss: 4.5012784539721906e-05
step: 110, loss: 0.0015269003342837095
step: 120, loss: 3.7507445085793734e-05
step: 130, loss: 2.8428996301954612e-05
step: 140, loss: 0.0003082027833443135
step: 150, loss: 0.0006562185590155423
step: 160, loss: 0.06210098788142204
step: 170, loss: 0.0073824478313326836
step: 180, loss: 0.0001251927751582116
step: 190, loss: 0.00015977023576851934
step: 200, loss: 0.0037479354068636894
step: 210, loss: 0.0013562508393079042
step: 220, loss: 0.00038754226989112794
step: 230, loss: 0.0002702068886719644
step: 240, loss: 9.359812247566879e-05
step: 250, loss: 4.0693212213227525e-05
step: 260, loss: 6.36223703622818e-05
step: 270, loss: 1.4114910300122574e-05
step: 280, loss: 8.423141116509214e-05
step: 290, loss: 1.6692465578671545e-05
step: 300, loss: 0.004016796126961708
step: 310, loss: 0.038503121584653854
step: 320, loss: 0.00524506438523531

```

```

step: 330, loss: 6.214376480784267e-05
step: 340, loss: 0.0005933194770477712
step: 350, loss: 0.00012064059410477057
step: 360, loss: 1.7433530956623144e-05
step: 370, loss: 1.869994775915984e-05
step: 380, loss: 0.0004731090448331088
step: 390, loss: 1.770171547832433e-05
step: 400, loss: 0.002418083604425192
step: 410, loss: 1.5340376194217242e-05
step: 420, loss: 3.259041113778949e-05
step: 430, loss: 8.702175364305731e-06
step: 440, loss: 1.4725616892974358e-05
step: 450, loss: 0.006922766100615263
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 460, loss: 0.00013307614426594228
step: 470, loss: 0.0016001507174223661
step: 480, loss: 0.00301870284602046
step: 490, loss: 0.00012905194307677448
step: 500, loss: 0.032723285257816315
step: 510, loss: 1.0166184438276105e-05
step: 520, loss: 0.00023470804444514215
step: 530, loss: 0.08995148539543152
/usr/local/lib/python3.7/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=0.9475164011246484, f1=0.9393090569561158,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.794383519561961e-05
step: 10, loss: 0.004440520890057087
step: 20, loss: 4.2193529225187376e-05
step: 30, loss: 5.519598926184699e-05
step: 40, loss: 0.00010854657739400864
step: 50, loss: 6.200841744430363e-05
step: 60, loss: 6.132567068561912e-05
step: 70, loss: 2.8754540835507214e-05
step: 80, loss: 6.007185220369138e-05
step: 90, loss: 0.001194715965539217
step: 100, loss: 0.0002920800179708749
step: 110, loss: 0.0005902857519686222
step: 120, loss: 2.6723777409642935e-05
step: 130, loss: 0.00026661239098757505
step: 140, loss: 5.142902955412865e-05
step: 150, loss: 0.0007430792902596295
step: 160, loss: 0.0005051809130236506
step: 170, loss: 0.0030917644035071135

```



```

step: 180, loss: 7.90422927821055e-05
step: 190, loss: 9.709729056339711e-05
step: 200, loss: 1.559008887852542e-05
step: 210, loss: 0.0003184875240549445
step: 220, loss: 0.006941680330783129
step: 230, loss: 3.016948357981164e-05
step: 240, loss: 0.00826103426516056
step: 250, loss: 0.00020329182734712958
step: 260, loss: 6.476557609857991e-05
step: 270, loss: 0.003703460330143571
step: 280, loss: 0.00016622785187792033
step: 290, loss: 1.3414536624622997e-05
step: 300, loss: 1.2226222679601051e-05
step: 310, loss: 0.00034295287332497537
step: 320, loss: 6.75428964314051e-05
step: 330, loss: 3.399378329049796e-05
step: 340, loss: 0.0041109537705779076
step: 350, loss: 2.978320662805345e-05
step: 360, loss: 0.00011553426884347573
step: 370, loss: 1.7824741007643752e-05
step: 380, loss: 8.845917182043195e-05
step: 390, loss: 3.7580146454274654e-05
step: 400, loss: 1.5530336895608343e-05
step: 410, loss: 0.0003705544222611934
step: 420, loss: 2.187629070249386e-05
step: 430, loss: 0.0047829775139689445
step: 440, loss: 2.2423795599024743e-05
step: 450, loss: 1.5071968846314121e-05
step: 460, loss: 1.8182521671405993e-05
step: 470, loss: 0.014541452750563622
step: 480, loss: 0.0001397078885929659
step: 490, loss: 0.0001096410778700374
step: 500, loss: 0.0012418117839843035
step: 510, loss: 3.326290607219562e-05
step: 520, loss: 2.464580575178843e-05
step: 530, loss: 0.00426693307235837
/usr/local/lib/python3.7/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=0.9480037140204272, f1=0.9382488479262673,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005078744143247604
step: 10, loss: 0.011167554184794426
step: 20, loss: 0.0007136341882869601
step: 30, loss: 2.3135051378631033e-05

```

step: 40, loss: 3.95783172280062e-05
step: 50, loss: 4.094407267984934e-05
step: 60, loss: 6.230139842955396e-05
step: 70, loss: 0.0005102158756926656
step: 80, loss: 0.0002881104010157287
step: 90, loss: 0.0003828128392342478
step: 100, loss: 0.03341246768832207
step: 110, loss: 3.787444802583195e-05
step: 120, loss: 9.491165110375732e-05
step: 130, loss: 0.00018896414258051664
step: 140, loss: 2.2991343939793296e-05
step: 150, loss: 0.0009021624573506415
step: 160, loss: 8.092899224720895e-05
step: 170, loss: 0.0002171232335967943
step: 180, loss: 8.481623808620498e-05
step: 190, loss: 0.0001360871974611655
step: 200, loss: 0.0013373655965551734
step: 210, loss: 2.8240105166332796e-05
step: 220, loss: 6.846587348263711e-05
step: 230, loss: 3.942261537304148e-05
step: 240, loss: 0.00011705914948834106
step: 250, loss: 2.6850233552977443e-05
step: 260, loss: 1.9310673451400362e-05
step: 270, loss: 0.0002956461685243994
step: 280, loss: 0.00011099260882474482
step: 290, loss: 3.727906005224213e-05
step: 300, loss: 0.0007492162985727191
step: 310, loss: 3.333613130962476e-05
step: 320, loss: 0.0013450148981064558
step: 330, loss: 0.0028469297103583813
step: 340, loss: 0.002188236452639103
step: 350, loss: 1.677355066931341e-05
step: 360, loss: 4.540313602774404e-05
step: 370, loss: 0.045717865228652954
step: 380, loss: 4.3655822082655504e-05
step: 390, loss: 0.005853912327438593
step: 400, loss: 5.218467049417086e-05
step: 410, loss: 0.0002781694056466222
step: 420, loss: 5.620465526590124e-05
step: 430, loss: 1.955692277988419e-05
step: 440, loss: 3.6129251384409145e-05
step: 450, loss: 0.00012322486145421863
step: 460, loss: 1.98696980078239e-05
step: 470, loss: 0.003183545544743538
step: 480, loss: 7.797698344802484e-05
step: 490, loss: 2.89690142381005e-05
step: 500, loss: 6.539450259879231e-05
step: 510, loss: 0.0007016057497821748

```

step: 520, loss: 3.554693830665201e-05
step: 530, loss: 8.347250695805997e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9484440315838365, f1=0.9400369003690037,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022411514073610306
step: 10, loss: 2.0253715774742886e-05
step: 20, loss: 0.0034920659381896257
step: 30, loss: 8.884740054782014e-06
step: 40, loss: 0.0001009205006994307
step: 50, loss: 0.0014828558778390288
step: 60, loss: 0.0029231372755020857
step: 70, loss: 3.521850157994777e-05
step: 80, loss: 0.00534677691757679
step: 90, loss: 6.439466960728168e-05
step: 100, loss: 1.939683897944633e-05
step: 110, loss: 0.0008841506787575781
step: 120, loss: 0.0003226099943276495
step: 130, loss: 1.5086502571648452e-05
step: 140, loss: 3.1250998290488496e-05
step: 150, loss: 1.2781320037902333e-05
step: 160, loss: 0.0001359682355541736
step: 170, loss: 2.5052790078916587e-05
step: 180, loss: 1.7068892702809535e-05
step: 190, loss: 2.4036995455389842e-05
step: 200, loss: 1.3298948942974675e-05
step: 210, loss: 1.7917875084094703e-05
step: 220, loss: 1.7046351786120795e-05
step: 230, loss: 0.001022382639348507
step: 240, loss: 8.811176667222753e-05
step: 250, loss: 2.3374799638986588e-05
step: 260, loss: 6.893922545714304e-05
step: 270, loss: 3.8675036194035783e-05
step: 280, loss: 5.2998024330008775e-05
step: 290, loss: 4.254471787135117e-05
step: 300, loss: 1.5269139112206176e-05
step: 310, loss: 0.0003072945983149111
step: 320, loss: 2.705209408304654e-05
step: 330, loss: 1.1470041499705985e-05
step: 340, loss: 4.4565764255821705e-05
step: 350, loss: 6.888001735205762e-06
step: 360, loss: 0.08770023286342621
step: 370, loss: 0.0001784299238352105

```

```

step: 380, loss: 1.5146734767768066e-05
step: 390, loss: 0.00015523449110332876
step: 400, loss: 2.276775376230944e-05
step: 410, loss: 1.2427381989255082e-05
step: 420, loss: 0.00020722548651974648
step: 430, loss: 3.797906538238749e-05
step: 440, loss: 2.609243347251322e-05
step: 450, loss: 1.4751620255992748e-05
step: 460, loss: 0.00027946539921686053
step: 470, loss: 0.0006826035096310079
step: 480, loss: 7.532466952397954e-06
step: 490, loss: 6.744228448951617e-05
step: 500, loss: 6.269154255278409e-05
step: 510, loss: 3.3709344279486686e-05
step: 520, loss: 1.5224301932903472e-05
step: 530, loss: 4.015819649794139e-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=0.9487895716945998, f1=0.9399260628465804,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.43114181002602e-05
step: 10, loss: 5.5410822824342176e-05
step: 20, loss: 3.34122269123327e-05
step: 30, loss: 0.00014603001181967556
step: 40, loss: 1.24866419355385e-05
step: 50, loss: 1.785478525562212e-05
step: 60, loss: 0.00015226777759380639
step: 70, loss: 5.619073999696411e-05
step: 80, loss: 0.0046277521178126335
step: 90, loss: 7.405820724670775e-06
step: 100, loss: 9.111980034504086e-06
step: 110, loss: 3.260956873418763e-05
step: 120, loss: 8.188118954421952e-06
step: 130, loss: 1.1495947546791285e-05
step: 140, loss: 6.000838038744405e-05
step: 150, loss: 1.2933870493725408e-05
step: 160, loss: 1.2688055903709028e-05
step: 170, loss: 1.392458852933487e-05
step: 180, loss: 0.00013371568638831377
step: 190, loss: 2.2327349142869934e-05
step: 200, loss: 1.9522494767443277e-05
step: 210, loss: 9.146754746325314e-05
step: 220, loss: 6.813505478930892e-06
step: 230, loss: 9.551503353577573e-06

```

```

step: 240, loss: 2.975189454446081e-05
step: 250, loss: 0.007727376185357571
step: 260, loss: 7.111534159776056e-06
step: 270, loss: 3.414889579289593e-05
step: 280, loss: 2.686602783796843e-05
step: 290, loss: 8.873434126144275e-06
step: 300, loss: 8.05401214165613e-06
step: 310, loss: 1.1980172530456912e-05
step: 320, loss: 7.4505232987576164e-06
step: 330, loss: 5.3967058192938566e-05
step: 340, loss: 3.0089075153227895e-05
step: 350, loss: 5.856130883330479e-06
step: 360, loss: 8.985288332041819e-06
step: 370, loss: 0.0002770027203951031
step: 380, loss: 0.000904237327631563
step: 390, loss: 0.0011260772589594126
step: 400, loss: 0.01694120652973652
step: 410, loss: 2.6330268156016245e-05
step: 420, loss: 6.761353233741829e-06
step: 430, loss: 3.640558134065941e-05
step: 440, loss: 0.0002144140744348988
step: 450, loss: 1.1659959454846103e-05
step: 460, loss: 9.624647645978257e-05
step: 470, loss: 6.40746247881907e-06
step: 480, loss: 1.0814254892466124e-05
step: 490, loss: 9.324304301117081e-06
step: 500, loss: 6.835861768195173e-06
step: 510, loss: 0.02831573598086834
step: 520, loss: 6.347825546981767e-06
step: 530, loss: 0.0019455758156254888
/usr/local/lib/python3.7/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=0.9507593189139439, f1=0.9397260273972603,
best_f1=0.9424326833797586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.372289474005811e-06
step: 10, loss: 1.1652303328446578e-05
step: 20, loss: 1.7868911527330056e-05
step: 30, loss: 0.0002689447719603777
step: 40, loss: 3.5080734960502014e-05
step: 50, loss: 0.005279161036014557
step: 60, loss: 1.2557607988128439e-05
step: 70, loss: 9.10807284526527e-06
step: 80, loss: 5.632615284412168e-06
step: 90, loss: 6.273348844842985e-06

```

```
step: 100, loss: 4.328772774897516e-06
step: 110, loss: 9.884744213195518e-05
step: 120, loss: 8.571774742449634e-06
step: 130, loss: 1.1715608707163483e-05
step: 140, loss: 8.292385246022604e-06
step: 150, loss: 7.610709417349426e-06
step: 160, loss: 6.373934866132913e-06
step: 170, loss: 5.785334451502422e-06
step: 180, loss: 4.903627632302232e-05
step: 190, loss: 9.273404430132359e-05
step: 200, loss: 1.8810103938449174e-05
step: 210, loss: 4.958337285643211e-06
step: 220, loss: 6.500587915070355e-06
step: 230, loss: 0.0001654895895626396
step: 240, loss: 2.9730479582212865e-05
step: 250, loss: 6.347849648591364e-06
step: 260, loss: 5.1297051868459675e-06
step: 270, loss: 9.350297659693751e-06
step: 280, loss: 7.469148386007873e-06
step: 290, loss: 0.0030256547033786774
step: 300, loss: 1.1104636541858781e-05
step: 310, loss: 0.014642040245234966
step: 320, loss: 8.575518222642131e-06
step: 330, loss: 7.357398772001034e-06
step: 340, loss: 3.409303826629184e-05
step: 350, loss: 0.0021838892716914415
step: 360, loss: 1.0710043170547578e-05
step: 370, loss: 9.450895959162153e-06
step: 380, loss: 6.45216414341121e-06
step: 390, loss: 1.5641460777260363e-05
step: 400, loss: 0.00034690325264818966
step: 410, loss: 1.1652224202407524e-05
step: 420, loss: 7.238184025482042e-06
step: 430, loss: 6.169022071844665e-06
step: 440, loss: 6.884295999043388e-06
step: 450, loss: 0.0003819829726126045
step: 460, loss: 8.928960596676916e-05
step: 470, loss: 5.822595539939357e-06
step: 480, loss: 0.0004500889335758984
step: 490, loss: 3.103598646703176e-05
step: 500, loss: 1.1052486115659121e-05
step: 510, loss: 5.8784717111848295e-06
step: 520, loss: 0.0024164849892258644
step: 530, loss: 7.882574209361337e-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=0.9526871841984381, f1=0.9390354868061875,
```

```
best_f1=0.9424326833797586
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
5742it [00:26, 214.57it/s]
load_f1 = 0.9548627268496975
real_f1 = 0.9512652296157451
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:24, 178.94it/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>, [247](https://us-python.pkg.dev/colab-</p></div><div data-bbox=)

```
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                                     | 827.9 MB 1.3 MB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: 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: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: srsly<1.1.0,>=0.1.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: 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: 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: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: 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: 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: 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: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: 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: 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: 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=4baf6012189777b40f11f4eda0a453fd1c900e4ca526e51a88791abf61eaaad0
  Stored in directory: /tmp/pip-ephem-wheel-cache-5dghoraw/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: 0.4455261826515198
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=0.27184466019417475, f1=0.32876712328767127,
best_f1=0.32876712328767127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44729769229888916
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.32876712328767127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41643595695495605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.29629629629629634, f1=0.13333333333333333,
best_f1=0.13333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23967649042606354

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.34285714285714286, f1=0.24390243902439024,
best_f1=0.24390243902439024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2773329019546509
/usr/local/lib/python3.7/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=0.34285714285714286, f1=0.2, best_f1=0.24390243902439024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20244988799095154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3636363636363636, f1=0.34374999999999994,
best_f1=0.34374999999999994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4755539298057556
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.47058823529411764, f1=0.44444444444444445,
best_f1=0.44444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48466625809669495

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6153846153846153, f1=0.56, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1930168867111206
/usr/local/lib/python3.7/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=0.6111111111111112, f1=0.6060606060606061, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.274048388004303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8148148148148148, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.278800904750824
/usr/local/lib/python3.7/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=0.717948717948718, f1=0.6857142857142857,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11993704736232758
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.8148148148148148,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07881737500429153
/usr/local/lib/python3.7/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=0.75, f1=0.75, best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08299315720796585
/usr/local/lib/python3.7/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=0.75, f1=0.75, best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12747129797935486
/usr/local/lib/python3.7/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=0.75, f1=0.75, best_f1=0.6666666666666666
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
91it [00:00, 117912.16it/s]
/usr/local/lib/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 = 0.6666666666666666
real_f1 = 0.6428571428571429
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 180.67it/s]

```

3.1.2 DBLP-ACM - Running the matcher

```

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

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

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

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.6038411855697632
```

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: 0.5301123261451721
```

```
step: 20, loss: 0.6014789938926697
```

```
step: 30, loss: 0.30411961674690247
```

```
step: 40, loss: 0.3494540750980377
```

```
step: 50, loss: 0.5497896671295166
```

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

```
step: 60, loss: 0.28429946303367615
```

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

```
step: 70, loss: 0.08185417205095291
```

```
step: 80, loss: 0.107674740254879
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0

```
step: 90, loss: 0.3745182752609253
```

```
step: 100, loss: 0.17719881236553192
```

```
step: 110, loss: 0.18002092838287354
```

```
step: 120, loss: 0.08473823219537735
```

```
step: 130, loss: 0.01753227226436138
```

```
step: 140, loss: 0.007738980930298567
```

```
step: 150, loss: 0.011298316530883312
```

```
step: 160, loss: 0.02510567381978035
```



```

step: 170, loss: 0.01909954659640789
step: 180, loss: 0.018365895375609398
step: 190, loss: 0.15328770875930786
step: 200, loss: 0.0867021307349205
step: 210, loss: 0.13015566766262054
step: 220, loss: 0.019597221165895462
step: 230, loss: 0.0025635443162173033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.972972972972973, f1=0.9774266365688488,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04358610138297081
step: 10, loss: 0.089680515229702
step: 20, loss: 0.022929158061742783
step: 30, loss: 0.01038315985351801
step: 40, loss: 0.030466875061392784
step: 50, loss: 0.007589625660330057
step: 60, loss: 0.009373089298605919
step: 70, loss: 0.016704903915524483
step: 80, loss: 0.005564062390476465
step: 90, loss: 0.0009619499323889613
step: 100, loss: 0.0034850446972995996
step: 110, loss: 0.028369758278131485
step: 120, loss: 0.010964998044073582
step: 130, loss: 0.004078483209013939
step: 140, loss: 0.002125371713191271
step: 150, loss: 0.12121588736772537
step: 160, loss: 0.0025350532960146666
step: 170, loss: 0.004536310210824013
step: 180, loss: 0.005794174503535032
step: 190, loss: 0.07629721611738205
step: 200, loss: 0.005066392477601767
step: 210, loss: 0.06630777567625046
step: 220, loss: 0.012047173455357552
step: 230, loss: 0.0004906051326543093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9740698985343857, f1=0.9728506787330317,
best_f1=0.9728506787330317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008596598170697689
step: 10, loss: 0.055974237620830536
step: 20, loss: 0.11630965769290924
step: 30, loss: 0.007309585344046354
step: 40, loss: 0.007133653853088617
step: 50, loss: 0.05791918560862541
step: 60, loss: 0.026545312255620956
step: 70, loss: 0.008017185144126415
step: 80, loss: 0.0008488500607199967
step: 90, loss: 0.004187884274870157
step: 100, loss: 0.0016344566829502583
step: 110, loss: 0.0018789914902299643
step: 120, loss: 0.0006771547487005591
step: 130, loss: 0.0015622101491317153
step: 140, loss: 0.001145332818850875
step: 150, loss: 0.018273087218403816
step: 160, loss: 0.008028673939406872
step: 170, loss: 0.0007157225627452135
step: 180, loss: 0.033637192100286484
step: 190, loss: 0.02557273581624031
step: 200, loss: 0.04070090502500534
step: 210, loss: 0.009982981719076633
step: 220, loss: 0.002694448223337531
step: 230, loss: 0.011391636915504932
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9863636363636363, f1=0.9759450171821306,
best_f1=0.9759450171821306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03972074016928673
step: 10, loss: 0.004017112776637077
step: 20, loss: 0.0009985683718696237
step: 30, loss: 0.00036581672611646354
step: 40, loss: 0.09453441202640533

```

```

step: 50, loss: 0.0017132731154561043
step: 60, loss: 0.0007656878442503512
step: 70, loss: 0.002896663034334779
step: 80, loss: 0.0007528535206802189
step: 90, loss: 0.003918540198355913
step: 100, loss: 0.00077995122410357
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 1024.0
step: 110, loss: 0.0011243289336562157
step: 120, loss: 0.08197766542434692
step: 130, loss: 0.0148170692846179
step: 140, loss: 0.01788390800356865
step: 150, loss: 0.009826863184571266
step: 160, loss: 0.08080146461725235
step: 170, loss: 0.016158320009708405
step: 180, loss: 0.11261770129203796
step: 190, loss: 0.011433849111199379
step: 200, loss: 0.027632759883999825
step: 210, loss: 0.018326958641409874
step: 220, loss: 0.009605174884200096
step: 230, loss: 0.0028083648066967726
/usr/local/lib/python3.7/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=0.9807909604519773, f1=0.9772209567198178,
best_f1=0.9759450171821306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016643961425870657
step: 10, loss: 0.005552122835069895
step: 20, loss: 0.0011312603019177914
step: 30, loss: 0.0028447872027754784
step: 40, loss: 0.0035580191761255264
step: 50, loss: 0.041008323431015015
step: 60, loss: 0.0021863970905542374
step: 70, loss: 0.00396115193143487
step: 80, loss: 0.08101394027471542
step: 90, loss: 0.08545149862766266
step: 100, loss: 0.0015613065334036946
step: 110, loss: 0.05453092232346535
step: 120, loss: 0.017782509326934814
step: 130, loss: 0.0013431069673970342
step: 140, loss: 0.00045358954230323434
step: 150, loss: 0.039901554584503174
step: 160, loss: 0.0007982976967468858
step: 170, loss: 0.030331986024975777
step: 180, loss: 0.00597818149253726
step: 190, loss: 0.013421095907688141

```

```

step: 200, loss: 0.0009661101503297687
step: 210, loss: 0.019105764105916023
step: 220, loss: 0.0009838031837716699
step: 230, loss: 0.025688879191875458
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987598647125141, f1=0.9830890642615557,
best_f1=0.9830890642615557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048443235573358834
step: 10, loss: 0.002376082818955183
step: 20, loss: 0.0014768503606319427
step: 30, loss: 0.0014398639323189855
step: 40, loss: 0.0002815349434968084
step: 50, loss: 0.00045228892122395337
step: 60, loss: 0.005222144071012735
step: 70, loss: 0.009486712515354156
step: 80, loss: 0.0017146425088867545
step: 90, loss: 0.003809885121881962
step: 100, loss: 0.0015525036724284291
step: 110, loss: 0.006912811193615198
step: 120, loss: 0.0003961448965128511
step: 130, loss: 0.001030834624543786
step: 140, loss: 0.001892576925456524
step: 150, loss: 0.0005835911724716425
step: 160, loss: 0.011153494007885456
step: 170, loss: 0.0004911596770398319
step: 180, loss: 0.03378280624747276
step: 190, loss: 0.00020021099771838635
step: 200, loss: 0.0006397637771442533
step: 210, loss: 0.0006013587699271739
step: 220, loss: 0.003928000573068857
step: 230, loss: 0.0026995260268449783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887892376681614, f1=0.9765886287625419,

```

```

best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002813709434121847
step: 10, loss: 0.0007939597126096487
step: 20, loss: 0.0029874658212065697
step: 30, loss: 0.0006070485105738044
step: 40, loss: 0.0006962298066355288
step: 50, loss: 0.001147167175076902
step: 60, loss: 0.014950500801205635
step: 70, loss: 0.0005547511391341686
step: 80, loss: 0.00461762398481369
step: 90, loss: 0.0019460961921140552
step: 100, loss: 0.00033740897197276354
step: 110, loss: 0.0006306427530944347
step: 120, loss: 0.004704300779849291
step: 130, loss: 0.007726042065769434
step: 140, loss: 0.0007859526085667312
step: 150, loss: 0.005797394551336765
step: 160, loss: 0.0005112368380650878
step: 170, loss: 0.00964822806417942
step: 180, loss: 0.00026070483727380633
step: 190, loss: 0.0002990987559314817
step: 200, loss: 0.053120147436857224
step: 210, loss: 0.049500782042741776
step: 220, loss: 0.0003748654271475971
step: 230, loss: 0.0017798530170693994
/usr/local/lib/python3.7/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=0.9842342342342343, f1=0.9774266365688488,
best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018715152982622385
step: 10, loss: 0.005974013824015856
step: 20, loss: 0.001494915341027081
step: 30, loss: 0.0015995721332728863
step: 40, loss: 0.00456612603738904
step: 50, loss: 0.0014252708060666919
step: 60, loss: 0.0007499870262108743
step: 70, loss: 0.0001862765639089048
step: 80, loss: 0.028395699337124825
step: 90, loss: 0.0009411269566044211
step: 100, loss: 0.0008635988342575729
step: 110, loss: 0.001729748211801052

```

```

step: 120, loss: 0.0026938433293253183
step: 130, loss: 0.0006972153205424547
step: 140, loss: 0.007799461018294096
step: 150, loss: 0.027800889685750008
step: 160, loss: 0.0154071981087327
step: 170, loss: 0.03440683335065842
step: 180, loss: 0.00040985149098560214
step: 190, loss: 0.0009168870747089386
step: 200, loss: 0.010413861833512783
step: 210, loss: 0.001662153284996748
step: 220, loss: 0.0009122680639848113
step: 230, loss: 0.00031728227622807026
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9787709497206705,
best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021377911616582423
step: 10, loss: 0.0008868036675266922
step: 20, loss: 0.0006400028942152858
step: 30, loss: 0.001038572401739657
step: 40, loss: 0.0005756777245551348
step: 50, loss: 0.0009292070171795785
step: 60, loss: 0.0005877428338862956
step: 70, loss: 0.023757144808769226
step: 80, loss: 0.0006365804583765566
step: 90, loss: 0.0310650784522295
step: 100, loss: 0.00030576030258089304
step: 110, loss: 0.0002195903070969507
step: 120, loss: 0.008298464119434357
step: 130, loss: 0.0031813844107091427
step: 140, loss: 0.0003339955292176455
step: 150, loss: 0.0005596216069534421
step: 160, loss: 0.0036277147009968758
step: 170, loss: 0.0013685954036191106
step: 180, loss: 0.0012165920343250036
step: 190, loss: 0.0002032312040682882
step: 200, loss: 0.00021372965420596302
step: 210, loss: 0.009862393140792847
step: 220, loss: 0.00016830085951369256
step: 230, loss: 0.001579958712682128
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9819819819819819,

```

```

best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020427012350410223
step: 10, loss: 0.0005356986657716334
step: 20, loss: 0.00032545262365601957
step: 30, loss: 9.569587564328685e-05
step: 40, loss: 0.0007737097912468016
step: 50, loss: 6.279148510657251e-05
step: 60, loss: 0.0001912511943373829
step: 70, loss: 0.0007670070044696331
step: 80, loss: 0.00016917353786993772
step: 90, loss: 0.000117877236334607
step: 100, loss: 0.00010241389099974185
step: 110, loss: 0.012753610499203205
step: 120, loss: 0.00011131393694086
step: 130, loss: 0.00026449200231581926
step: 140, loss: 0.00046796276001259685
step: 150, loss: 0.0008762399666011333
step: 160, loss: 7.142400136217475e-05
step: 170, loss: 0.00016268545005004853
step: 180, loss: 0.0008334098965860903
step: 190, loss: 0.00023402941587846726
step: 200, loss: 0.000676424300763756
step: 210, loss: 0.032044630497694016
step: 220, loss: 7.812853436917067e-05
step: 230, loss: 8.219205483328551e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9887387387387387, f1=0.9819004524886877,
best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.886570140020922e-05
step: 10, loss: 0.00014068960445001721
step: 20, loss: 0.0001243615843122825
step: 30, loss: 0.00010296619439031929
step: 40, loss: 2.655868956935592e-05
step: 50, loss: 6.28834095550701e-05
step: 60, loss: 0.009896083734929562
step: 70, loss: 6.198789196787402e-05
step: 80, loss: 0.0008812181185930967
step: 90, loss: 0.04242029786109924
step: 100, loss: 0.002356862183660269
step: 110, loss: 0.0007545602857135236

```

```

step: 120, loss: 0.00021564806229434907
step: 130, loss: 0.00013822065375279635
step: 140, loss: 0.0002749140840023756
step: 150, loss: 0.0001289007777813822
step: 160, loss: 0.07851088792085648
step: 170, loss: 0.0009854048257693648
step: 180, loss: 0.000642909377347678
step: 190, loss: 0.0001732706732582301
step: 200, loss: 0.0029363995417952538
step: 210, loss: 0.0005510759074240923
step: 220, loss: 0.0018644160591065884
step: 230, loss: 0.00020333443535491824
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9807909604519773,
best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018969419761560857
step: 10, loss: 5.981950380373746e-05
step: 20, loss: 0.01978805847465992
step: 30, loss: 0.02950344793498516
step: 40, loss: 9.950847743311897e-05
step: 50, loss: 0.0001068541023414582
step: 60, loss: 0.00019783001334872097
step: 70, loss: 6.496915739262477e-05
step: 80, loss: 3.709707743837498e-05
step: 90, loss: 0.0005180159932933748
step: 100, loss: 7.890518463682383e-05
step: 110, loss: 5.467044684337452e-05
step: 120, loss: 9.164876973954961e-05
step: 130, loss: 6.412194488802925e-05
step: 140, loss: 0.00010294420644640923
step: 150, loss: 7.894742157077417e-05
step: 160, loss: 0.00037176060141064227
step: 170, loss: 6.218014459591359e-05
step: 180, loss: 5.7385070249438286e-05
step: 190, loss: 0.011286252178251743
step: 200, loss: 9.539545135339722e-05
step: 210, loss: 0.0002779607893899083
step: 220, loss: 0.020872173830866814
step: 230, loss: 0.00040469097439199686
/usr/local/lib/python3.7/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=0.9886621315192743, f1=0.9829351535836178,

```



```

best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.058662256691605e-05
step: 10, loss: 6.296303035924211e-05
step: 20, loss: 8.247750520240515e-05
step: 30, loss: 2.8827971618738957e-05
step: 40, loss: 6.683432002319023e-05
step: 50, loss: 0.00159190664999187
step: 60, loss: 5.682041228283197e-05
step: 70, loss: 3.854671012959443e-05
step: 80, loss: 3.2094816560857e-05
step: 90, loss: 4.297896521165967e-05
step: 100, loss: 8.73986937222071e-05
step: 110, loss: 0.0002532113285269588
step: 120, loss: 8.478414383716881e-05
step: 130, loss: 6.390047929016873e-05
step: 140, loss: 4.86919452669099e-05
step: 150, loss: 2.398214019194711e-05
step: 160, loss: 0.026197075843811035
step: 170, loss: 5.9036297898273915e-05
step: 180, loss: 0.036377862095832825
step: 190, loss: 6.698729703202844e-05
step: 200, loss: 3.456833655945957e-05
step: 210, loss: 0.0001902063231682405
step: 220, loss: 5.171047450858168e-05
step: 230, loss: 0.00011404855467844754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9887892376681614, f1=0.9808342728297633,
best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.047962571959943e-05
step: 10, loss: 5.114505620440468e-05
step: 20, loss: 7.163298869272694e-05
step: 30, loss: 8.822892414173111e-05
step: 40, loss: 4.682708822656423e-05
step: 50, loss: 5.33322490809951e-05
step: 60, loss: 5.189944204175845e-05
step: 70, loss: 4.3359133997000754e-05
step: 80, loss: 3.5179255064576864e-05
step: 90, loss: 7.092089072102681e-05
step: 100, loss: 4.426038867677562e-05
step: 110, loss: 5.532054274226539e-05

```

```

step: 120, loss: 2.0704304915852845e-05
step: 130, loss: 6.41403385088779e-05
step: 140, loss: 3.7909620004938915e-05
step: 150, loss: 3.136470331810415e-05
step: 160, loss: 4.9999740440398455e-05
step: 170, loss: 5.579539720201865e-05
step: 180, loss: 4.43128046754282e-05
step: 190, loss: 0.0003559836186468601
step: 200, loss: 9.639786730986089e-05
step: 210, loss: 0.00010876672604354098
step: 220, loss: 4.322913082432933e-05
step: 230, loss: 2.5990182621171698e-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=0.9865771812080537, f1=0.9775784753363228,
best_f1=0.9765886287625419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015217810869216919
step: 10, loss: 3.9526548789581284e-05
step: 20, loss: 6.857956759631634e-05
step: 30, loss: 3.643138916231692e-05
step: 40, loss: 2.7800442694569938e-05
step: 50, loss: 3.614788511185907e-05
step: 60, loss: 0.025640057399868965
step: 70, loss: 3.245680272812024e-05
step: 80, loss: 3.5663866583490744e-05
step: 90, loss: 3.3421350963180885e-05
step: 100, loss: 2.8794898753403686e-05
step: 110, loss: 0.0002152538945665583
step: 120, loss: 0.02230699174106121
step: 130, loss: 2.7319958462612703e-05
step: 140, loss: 0.024392640218138695
step: 150, loss: 0.000124288271763362
step: 160, loss: 4.579054075293243e-05
step: 170, loss: 2.7118518119095825e-05
step: 180, loss: 5.0268346967641264e-05
step: 190, loss: 9.768297604750842e-05
step: 200, loss: 0.00015084221377037466
step: 210, loss: 0.04562801867723465
step: 220, loss: 4.3235981138423085e-05
step: 230, loss: 5.311031418386847e-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=0.9876819708846584, f1=0.9786276715410572,

```

```

best_f1=0.9765886287625419
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, 180.09it/s]
load_f1 = 0.9887892376681614
real_f1 = 0.9854748603351955
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, 167.93it/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: 0.6633002161979675
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: 0.4148438274860382
step: 20, loss: 0.2740328907966614
step: 30, loss: 0.34824463725090027
step: 40, loss: 0.3712451159954071
step: 50, loss: 0.5943734049797058
step: 60, loss: 0.32668256759643555
step: 70, loss: 0.4528818726539612
step: 80, loss: 0.28014758229255676
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.3125389516353607
step: 100, loss: 0.28614190220832825
step: 110, loss: 0.16533403098583221
step: 120, loss: 0.23268966376781464
step: 130, loss: 0.22107017040252686
step: 140, loss: 0.24529296159744263
step: 150, loss: 0.10652834177017212
step: 160, loss: 0.4318292737007141
step: 170, loss: 0.24340985715389252
step: 180, loss: 0.17252673208713531
step: 190, loss: 0.1444459855556488
step: 200, loss: 0.027625590562820435
step: 210, loss: 0.06703818589448929
step: 220, loss: 0.21251754462718964
step: 230, loss: 0.15305893123149872
step: 240, loss: 0.03334831818938255
step: 250, loss: 0.16459417343139648
step: 260, loss: 0.13200300931930542
step: 270, loss: 0.5005769729614258
step: 280, loss: 0.07783783972263336
step: 290, loss: 0.08746127784252167
step: 300, loss: 0.06664042174816132
step: 310, loss: 0.10136877000331879
step: 320, loss: 0.05649387463927269
step: 330, loss: 0.110294409096241
step: 340, loss: 0.34135305881500244
step: 350, loss: 0.16996128857135773
step: 360, loss: 0.07086192816495895
step: 370, loss: 0.08073178678750992
step: 380, loss: 0.2821148931980133

```

```

step: 390, loss: 0.05135170370340347
step: 400, loss: 0.041219860315322876
step: 410, loss: 0.36446523666381836
step: 420, loss: 0.047029171139001846
step: 430, loss: 0.046573080122470856
step: 440, loss: 0.05256876349449158
step: 450, loss: 0.13140837848186493
step: 460, loss: 0.12140364199876785
step: 470, loss: 0.04102642089128494
step: 480, loss: 0.157316654920578
step: 490, loss: 0.13170652091503143
step: 500, loss: 0.040289971977472305
step: 510, loss: 0.05028347298502922
step: 520, loss: 0.07637792825698853
step: 530, loss: 0.22066886723041534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9242009132420091, f1=0.9245454545454544,
best_f1=0.9245454545454544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08949863165616989
step: 10, loss: 0.055857449769973755
step: 20, loss: 0.04780464246869087
step: 30, loss: 0.04032830148935318
step: 40, loss: 0.07635578513145447
step: 50, loss: 0.07894771546125412
step: 60, loss: 0.049136653542518616
step: 70, loss: 0.12516506016254425
step: 80, loss: 0.0965457558631897
step: 90, loss: 0.03204255551099777
step: 100, loss: 0.18007643520832062
step: 110, loss: 0.012471344321966171
step: 120, loss: 0.061274368315935135
step: 130, loss: 0.049883272498846054
step: 140, loss: 0.11205585300922394
step: 150, loss: 0.03700680658221245
step: 160, loss: 0.049664195626974106
step: 170, loss: 0.023966170847415924
step: 180, loss: 0.018221024423837662
step: 190, loss: 0.041271861642599106
step: 200, loss: 0.2161022573709488

```

```

step: 210, loss: 0.027680503204464912
step: 220, loss: 0.002421043114736676
step: 230, loss: 0.041640955954790115
step: 240, loss: 0.023105962201952934
step: 250, loss: 0.04812060669064522
step: 260, loss: 0.107023686170578
step: 270, loss: 0.08186066895723343
step: 280, loss: 0.04170987755060196
step: 290, loss: 0.016361473128199577
step: 300, loss: 0.023874381557106972
step: 310, loss: 0.028567012399435043
step: 320, loss: 0.06116918474435806
step: 330, loss: 0.10188967734575272
step: 340, loss: 0.06922028213739395
step: 350, loss: 0.018361438065767288
step: 360, loss: 0.059423040598630905
step: 370, loss: 0.04467778280377388
step: 380, loss: 0.11413725465536118
step: 390, loss: 0.01401477586477995
step: 400, loss: 0.02519138529896736
step: 410, loss: 0.043841708451509476
step: 420, loss: 0.15103694796562195
step: 430, loss: 0.16211281716823578
step: 440, loss: 0.0071416194550693035
step: 450, loss: 0.0496302954852581
step: 460, loss: 0.057666487991809845
step: 470, loss: 0.01958024501800537
step: 480, loss: 0.012840621173381805
step: 490, loss: 0.10608899593353271
step: 500, loss: 0.030106021091341972
step: 510, loss: 0.03706321120262146
step: 520, loss: 0.380759060382843
step: 530, loss: 0.16642238199710846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9459962756052142, f1=0.9412861136999067,
best_f1=0.9412861136999067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10141416639089584
step: 10, loss: 0.029501279816031456
step: 20, loss: 0.010505826212465763

```

step: 30, loss: 0.08849829435348511
step: 40, loss: 0.056656595319509506
step: 50, loss: 0.029805276542901993
step: 60, loss: 0.057562824338674545
step: 70, loss: 0.004698685370385647
step: 80, loss: 0.01716931164264679
step: 90, loss: 0.032052118331193924
step: 100, loss: 0.042644258588552475
step: 110, loss: 0.041154537349939346
step: 120, loss: 0.20637866854667664
step: 130, loss: 0.055472951382398605
step: 140, loss: 0.01623453199863434
step: 150, loss: 0.038815416395664215
step: 160, loss: 0.011357192881405354
step: 170, loss: 0.012323918752372265
step: 180, loss: 0.02295919694006443
step: 190, loss: 0.014086339622735977
step: 200, loss: 0.02605254389345646
step: 210, loss: 0.03477204591035843
step: 220, loss: 0.09323188662528992
step: 230, loss: 0.04963100329041481
step: 240, loss: 0.10949353128671646
step: 250, loss: 0.20394931733608246
step: 260, loss: 0.10624699294567108
step: 270, loss: 0.02867463417351246
step: 280, loss: 0.06135249882936478
step: 290, loss: 0.12668099999427795
step: 300, loss: 0.11801011860370636
step: 310, loss: 0.09548146277666092
step: 320, loss: 0.046513646841049194
step: 330, loss: 0.0074676028452813625
step: 340, loss: 0.014503000304102898
step: 350, loss: 0.09589576721191406
step: 360, loss: 0.009421447291970253
step: 370, loss: 0.09232217818498611
step: 380, loss: 0.018594952300190926
step: 390, loss: 0.00419800728559494
step: 400, loss: 0.07247129827737808
step: 410, loss: 0.02592460811138153
step: 420, loss: 0.020667267963290215
step: 430, loss: 0.027538076043128967
step: 440, loss: 0.22676828503608704
step: 450, loss: 0.01134430430829525
step: 460, loss: 0.0641363263130188
step: 470, loss: 0.01061617024242878
step: 480, loss: 0.09748475253582001
step: 490, loss: 0.0050543323159217834
step: 500, loss: 0.04013563692569733


```

step: 510, loss: 0.08667320758104324
step: 520, loss: 0.016191110014915466
step: 530, loss: 0.005028291139751673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9475638051044084, f1=0.9407407407407408,
best_f1=0.9407407407407408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051063355058431625
step: 10, loss: 0.00392047269269824
step: 20, loss: 0.174769327044487
step: 30, loss: 0.09162045270204544
step: 40, loss: 0.06841558963060379
step: 50, loss: 0.1054617166519165
step: 60, loss: 0.003989634104073048
step: 70, loss: 0.019989101216197014
step: 80, loss: 0.11267782747745514
step: 90, loss: 0.10011187195777893
step: 100, loss: 0.0012275835033506155
step: 110, loss: 0.04466306045651436
step: 120, loss: 0.001739670056849718
step: 130, loss: 0.06533178687095642
step: 140, loss: 0.043475642800331116
step: 150, loss: 0.002968843560665846
step: 160, loss: 0.029862895607948303
step: 170, loss: 0.030103206634521484
step: 180, loss: 0.0387626513838768
step: 190, loss: 0.07137357443571091
step: 200, loss: 0.10470839589834213
step: 210, loss: 0.005065463948994875
step: 220, loss: 0.1302160769701004
step: 230, loss: 0.004015476442873478
step: 240, loss: 0.0035426500253379345
step: 250, loss: 0.18818065524101257
step: 260, loss: 0.0017500227550044656
step: 270, loss: 0.09140487760305405
step: 280, loss: 0.012478198856115341
step: 290, loss: 0.07081977277994156
step: 300, loss: 0.02349177934229374
step: 310, loss: 0.0048771449364721775
step: 320, loss: 0.16081944108009338

```

```

step: 330, loss: 0.017105545848608017
step: 340, loss: 0.0020675405394285917
step: 350, loss: 0.24486367404460907
step: 360, loss: 0.05228249728679657
step: 370, loss: 0.007409068290144205
step: 380, loss: 0.007212957367300987
step: 390, loss: 0.003638830268755555
step: 400, loss: 0.017117232084274292
step: 410, loss: 0.009843673557043076
step: 420, loss: 0.003442309098318219
step: 430, loss: 0.006785341538488865
step: 440, loss: 0.0031444188207387924
step: 450, loss: 0.0037661301903426647
step: 460, loss: 0.04609248414635658
step: 470, loss: 0.021742699667811394
step: 480, loss: 0.007332568988204002
step: 490, loss: 0.011863185092806816
step: 500, loss: 0.014603016898036003
step: 510, loss: 0.2020021229982376
step: 520, loss: 0.008481043390929699
step: 530, loss: 0.06720175594091415
/usr/local/lib/python3.7/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=0.9347826086956521, f1=0.9352720450281427,
best_f1=0.9407407407407408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00345848617143929
step: 10, loss: 0.015576564706861973
step: 20, loss: 0.0026730732060968876
step: 30, loss: 0.03444598242640495
step: 40, loss: 0.0047210403718054295
step: 50, loss: 0.1525445431470871
step: 60, loss: 0.09606759995222092
step: 70, loss: 0.004166719503700733
step: 80, loss: 0.0016459355829283595
step: 90, loss: 0.008116289041936398
step: 100, loss: 0.006884741596877575
step: 110, loss: 0.02853008732199669
step: 120, loss: 0.1340763121843338
step: 130, loss: 0.015573451295495033
step: 140, loss: 0.02809293381869793
step: 150, loss: 0.011937426403164864
step: 160, loss: 0.004329086281359196
step: 170, loss: 0.09987609833478928
step: 180, loss: 0.012094035744667053

```

```

step: 190, loss: 0.002181239193305373
step: 200, loss: 0.012692553922533989
step: 210, loss: 0.006559773813933134
step: 220, loss: 0.006166442763060331
step: 230, loss: 0.007490357384085655
step: 240, loss: 0.0021318630315363407
step: 250, loss: 0.10241948068141937
step: 260, loss: 0.0030514297541230917
step: 270, loss: 0.002162665594369173
step: 280, loss: 0.009775730781257153
step: 290, loss: 0.03849324584007263
step: 300, loss: 0.10218816250562668
step: 310, loss: 0.058354027569293976
step: 320, loss: 0.0248107872903347
step: 330, loss: 0.0007387064397335052
step: 340, loss: 0.017529375851154327
step: 350, loss: 0.011567327193915844
step: 360, loss: 0.0009648058330640197
step: 370, loss: 0.002088970970362425
step: 380, loss: 0.00339457206428051
step: 390, loss: 0.02492385357618332
step: 400, loss: 0.012181345373392105
step: 410, loss: 0.07798461616039276
step: 420, loss: 0.1924920380115509
step: 430, loss: 0.0791456550359726
step: 440, loss: 0.0011186142219230533
step: 450, loss: 0.015625586733222008
step: 460, loss: 0.004163761157542467
step: 470, loss: 0.009213329292833805
step: 480, loss: 0.0018323963740840554
step: 490, loss: 0.0061677927151322365
step: 500, loss: 0.06218675523996353
step: 510, loss: 0.005562108475714922
step: 520, loss: 0.09015243500471115
step: 530, loss: 0.007972460240125656
/usr/local/lib/python3.7/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=0.9436947417403444, f1=0.9420491423273065,
best_f1=0.9407407407407408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035181406419724226
step: 10, loss: 0.0010742545127868652
step: 20, loss: 0.0015767146833240986
step: 30, loss: 0.002117904368788004
step: 40, loss: 0.037603817880153656

```

step: 50, loss: 0.0004237226094119251
step: 60, loss: 0.001563016208820045
step: 70, loss: 0.017232853919267654
step: 80, loss: 0.00017938134260475636
step: 90, loss: 0.056798674166202545
step: 100, loss: 0.0010517328046262264
step: 110, loss: 0.0022451584227383137
step: 120, loss: 0.0043974947184324265
step: 130, loss: 0.014241470023989677
step: 140, loss: 0.002595874946564436
step: 150, loss: 0.0022599902004003525
step: 160, loss: 0.007076721638441086
step: 170, loss: 0.001357779954560101
step: 180, loss: 0.003915566951036453
step: 190, loss: 0.035121217370033264
step: 200, loss: 0.027078501880168915
step: 210, loss: 0.0014767477987334132
step: 220, loss: 0.005164914298802614
step: 230, loss: 0.0011115085799247026
step: 240, loss: 0.0006016873521730304
step: 250, loss: 0.14038372039794922
step: 260, loss: 0.00020592391956597567
step: 270, loss: 0.0015049122739583254
step: 280, loss: 0.001230369322001934
step: 290, loss: 0.0011199346045032144
step: 300, loss: 0.007621431723237038
step: 310, loss: 0.07843036949634552
step: 320, loss: 0.0005194679251872003
step: 330, loss: 0.003319766139611602
step: 340, loss: 0.0009772221092134714
step: 350, loss: 0.003225277177989483
step: 360, loss: 0.00606259610503912
step: 370, loss: 0.003768308786675334
step: 380, loss: 0.001842169906012714
step: 390, loss: 0.004791955929249525
step: 400, loss: 0.0032524950802326202
step: 410, loss: 0.0009972768602892756
step: 420, loss: 0.0029691578820347786
step: 430, loss: 0.002742951735854149
step: 440, loss: 0.02984588034451008
step: 450, loss: 0.18628087639808655
step: 460, loss: 0.001452928176149726
step: 470, loss: 0.0017512327758595347
step: 480, loss: 0.009689467959105968
step: 490, loss: 0.017411191016435623
step: 500, loss: 0.004648905247449875
step: 510, loss: 0.08949360251426697
step: 520, loss: 0.0026903704274445772

```
step: 530, loss: 0.005595784168690443
/usr/local/lib/python3.7/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=0.94357515410147, f1=0.9408424041646947,
best_f1=0.9407407407407408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005565283354371786
step: 10, loss: 0.0016898594330996275
step: 20, loss: 0.005012927111238241
step: 30, loss: 0.02338939905166626
step: 40, loss: 0.08841172605752945
step: 50, loss: 0.013403872027993202
step: 60, loss: 0.003830250585451722
step: 70, loss: 0.0012486891355365515
step: 80, loss: 0.0004424243816174567
step: 90, loss: 0.0003363527648616582
step: 100, loss: 0.04327262192964554
step: 110, loss: 0.00014276517322286963
step: 120, loss: 0.0002838599029928446
step: 130, loss: 0.0007002754718996584
step: 140, loss: 0.0005459652747958899
step: 150, loss: 0.0007154330960474908
step: 160, loss: 0.0040311310440301895
step: 170, loss: 0.00031638069776818156
step: 180, loss: 0.0005326357204467058
step: 190, loss: 0.023459644988179207
step: 200, loss: 0.00015505115152336657
step: 210, loss: 0.001470611896365881
step: 220, loss: 0.0033659012988209724
step: 230, loss: 0.0035306329373270273
step: 240, loss: 0.00601207185536623
step: 250, loss: 0.0014615264954045415
step: 260, loss: 0.00018857588293030858
step: 270, loss: 0.00016703159781172872
step: 280, loss: 0.02190690115094185
step: 290, loss: 0.0020987954922020435
step: 300, loss: 0.0007223711581900716
step: 310, loss: 0.0012580036418512464
step: 320, loss: 0.0003087654185947031
step: 330, loss: 0.003336156252771616
step: 340, loss: 0.0004732045345008373
step: 350, loss: 0.0014132140204310417
step: 360, loss: 0.07212752848863602
step: 370, loss: 0.07843141257762909
step: 380, loss: 0.012461748905479908
```

```

step: 390, loss: 0.004683241713792086
step: 400, loss: 0.1519060879945755
step: 410, loss: 0.004558388609439135
step: 420, loss: 0.012035132385790348
step: 430, loss: 0.0028921174816787243
step: 440, loss: 0.0005927865859121084
step: 450, loss: 0.0036579580046236515
step: 460, loss: 0.0044378796592354774
step: 470, loss: 0.16118201613426208
step: 480, loss: 0.009555208496749401
step: 490, loss: 0.0014718043385073543
step: 500, loss: 0.004427107982337475
step: 510, loss: 0.0007734039681963623
step: 520, loss: 0.000837378844153136
step: 530, loss: 0.004193959292024374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9508497932935233, f1=0.9457720588235294,
best_f1=0.9457720588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005718591273762286
step: 10, loss: 0.0006227887351997197
step: 20, loss: 0.0004749142681248486
step: 30, loss: 0.00018997212464455515
step: 40, loss: 0.00010104570537805557
step: 50, loss: 0.0032329782843589783
step: 60, loss: 0.0003179177874699235
step: 70, loss: 0.0003768336900975555
step: 80, loss: 0.00027335959021002054
step: 90, loss: 0.00018108125368598849
step: 100, loss: 7.249003101605922e-05
step: 110, loss: 0.0001874118606792763
step: 120, loss: 0.00034126886748708785
step: 130, loss: 0.0128393629565835
step: 140, loss: 0.00045726820826530457
step: 150, loss: 9.851998038357124e-05
step: 160, loss: 0.00018576468573883176
step: 170, loss: 0.046619266271591187
step: 180, loss: 0.000329270405927673
step: 190, loss: 0.0011765939416363835
step: 200, loss: 0.0027124546468257904

```

```

step: 210, loss: 0.038732416927814484
step: 220, loss: 0.00014921546971891075
step: 230, loss: 0.082218237221241
step: 240, loss: 0.0005860731471329927
step: 250, loss: 0.0008017407380975783
step: 260, loss: 0.0005088448524475098
step: 270, loss: 0.005281486082822084
step: 280, loss: 0.0003179551858920604
step: 290, loss: 0.0016496900934726
step: 300, loss: 0.00013123768439982086
step: 310, loss: 0.0017792544094845653
step: 320, loss: 0.00020789091649930924
step: 330, loss: 0.0003673468600027263
step: 340, loss: 0.03601280599832535
step: 350, loss: 7.506808469770476e-05
step: 360, loss: 0.07981230318546295
step: 370, loss: 0.004620030056685209
step: 380, loss: 0.0018395379884168506
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 390, loss: 0.0022296514362096786
step: 400, loss: 0.013984136283397675
step: 410, loss: 0.0010743607999756932
step: 420, loss: 0.0007012312998995185
step: 430, loss: 0.0014684763737022877
step: 440, loss: 0.04536813870072365
step: 450, loss: 0.0011643313337117434
step: 460, loss: 0.0020510887261480093
step: 470, loss: 0.09835527092218399
step: 480, loss: 0.03837912902235985
step: 490, loss: 0.00031395984115079045
step: 500, loss: 0.00011043033009627834
step: 510, loss: 0.005708071403205395
step: 520, loss: 0.00028662860859185457
step: 530, loss: 0.06863473355770111
/usr/local/lib/python3.7/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=0.9479022591055786, f1=0.9404157043879908,
best_f1=0.9457720588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020988118194509298
step: 10, loss: 0.0030445209704339504
step: 20, loss: 0.0023620896972715855
step: 30, loss: 0.04578807204961777
step: 40, loss: 0.004818170331418514
step: 50, loss: 0.0012112336698919535

```

step: 60, loss: 0.0018608501413837075
step: 70, loss: 0.07366760820150375
step: 80, loss: 0.03945314884185791
step: 90, loss: 0.05018068850040436
step: 100, loss: 0.01215367577970028
step: 110, loss: 0.009688557125627995
step: 120, loss: 0.0008930990588851273
step: 130, loss: 0.0006724462728016078
step: 140, loss: 0.0003542932972777635
step: 150, loss: 0.023562144488096237
step: 160, loss: 0.018061628565192223
step: 170, loss: 0.002690351102501154
step: 180, loss: 0.0002086600143229589
step: 190, loss: 0.0003448408388067037
step: 200, loss: 0.0012035410618409514
step: 210, loss: 0.13170741498470306
step: 220, loss: 0.09910598397254944
step: 230, loss: 0.0005513360374607146
step: 240, loss: 0.000697821902576834
step: 250, loss: 0.00043977523455396295
step: 260, loss: 0.0009821628918871284
step: 270, loss: 0.024636128917336464
step: 280, loss: 0.0012242100201547146
step: 290, loss: 0.00023398928169626743
step: 300, loss: 0.0004264888702891767
step: 310, loss: 0.0619385801255703
step: 320, loss: 0.0016748050693422556
step: 330, loss: 0.000272925419267267
step: 340, loss: 0.003059764625504613
step: 350, loss: 0.1408718377351761
step: 360, loss: 0.0004402172635309398
step: 370, loss: 0.00022663341951556504
step: 380, loss: 0.0036313964519649744
step: 390, loss: 0.00012221832002978772
step: 400, loss: 0.04233304411172867
step: 410, loss: 0.00022955001622904092
step: 420, loss: 0.00017952480993699282
step: 430, loss: 0.0014065635623410344
step: 440, loss: 0.0010270995553582907
step: 450, loss: 0.0013716361718252301
step: 460, loss: 0.0003304299316368997
step: 470, loss: 0.0003013382665812969
step: 480, loss: 6.896359991515055e-05
step: 490, loss: 0.04900793731212616
step: 500, loss: 0.00020475378551054746
step: 510, loss: 0.023879630491137505
step: 520, loss: 0.0016411221586167812
step: 530, loss: 0.012652105651795864


```
/usr/local/lib/python3.7/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=0.9499536607970344, f1=0.9453015427769986,
best_f1=0.9457720588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016966609982773662
step: 10, loss: 0.0023294396232813597
step: 20, loss: 0.0012008537305518985
step: 30, loss: 0.005379924550652504
step: 40, loss: 0.0001651729253353551
step: 50, loss: 7.052782893879339e-05
step: 60, loss: 0.00018740609812084585
step: 70, loss: 0.00011448313307482749
step: 80, loss: 4.793186963070184e-05
step: 90, loss: 0.0006442955927923322
step: 100, loss: 0.003082016948610544
step: 110, loss: 0.0031744525767862797
step: 120, loss: 9.722570393932983e-05
step: 130, loss: 0.0013826315989717841
step: 140, loss: 0.001794298063032329
step: 150, loss: 0.0003803647996392101
step: 160, loss: 0.021856792271137238
step: 170, loss: 0.0001520083169452846
step: 180, loss: 0.000347673223586753
step: 190, loss: 0.00021503861353266984
step: 200, loss: 0.0482378713786602
step: 210, loss: 0.0004203195567242801
step: 220, loss: 0.00029568205354735255
step: 230, loss: 0.00020368538389448076
step: 240, loss: 0.00023789727129042149
step: 250, loss: 0.0004327953211031854
step: 260, loss: 0.0003434538666624576
step: 270, loss: 0.0002056693920167163
step: 280, loss: 0.011412782594561577
step: 290, loss: 0.0009536782163195312
step: 300, loss: 0.000504643889144063
step: 310, loss: 0.01841694861650467
step: 320, loss: 0.00249294051900506
step: 330, loss: 0.0009721522219479084
step: 340, loss: 8.737675671000034e-05
step: 350, loss: 0.003432716941460967
step: 360, loss: 0.0008463856065645814
step: 370, loss: 0.0005343094817362726
step: 380, loss: 0.001148969866335392
step: 390, loss: 0.0004110117442905903
```

```

step: 400, loss: 0.0002474125358276069
step: 410, loss: 8.661361789563671e-05
step: 420, loss: 7.997419743333012e-05
step: 430, loss: 5.829726796946488e-05
step: 440, loss: 8.99929873412475e-05
step: 450, loss: 0.00021343972184695303
step: 460, loss: 8.350752614205703e-05
step: 470, loss: 0.01216052658855915
step: 480, loss: 0.001236757612787187
step: 490, loss: 0.024200551211833954
step: 500, loss: 0.0083750756457448
step: 510, loss: 5.9512854932108894e-05
step: 520, loss: 0.00011722042108885944
step: 530, loss: 0.011355127207934856
/usr/local/lib/python3.7/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=0.9502304147465437, f1=0.9480037140204272,
best_f1=0.9457720588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022814172552898526
step: 10, loss: 0.00012421673454809934
step: 20, loss: 0.00039950941572897136
step: 30, loss: 0.0004812834376934916
step: 40, loss: 0.00023612129734829068
step: 50, loss: 0.00017754676810000092
step: 60, loss: 0.0001286200131289661
step: 70, loss: 0.0005188880022615194
step: 80, loss: 0.0002874484343919903
step: 90, loss: 0.0004545847768895328
step: 100, loss: 0.00046790274791419506
step: 110, loss: 0.0004647033056244254
step: 120, loss: 7.298964919755235e-05
step: 130, loss: 4.1395993321202695e-05
step: 140, loss: 0.000153832821524702
step: 150, loss: 0.0003549831744749099
step: 160, loss: 0.00022628255828749388
step: 170, loss: 0.00015544742927886546
step: 180, loss: 0.0002973258087877184
step: 190, loss: 0.0007639571558684111
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 200, loss: 9.710261656437069e-05
step: 210, loss: 0.0003838048141915351
step: 220, loss: 0.017385011538863182
step: 230, loss: 3.26414083247073e-05
step: 240, loss: 0.011531814001500607

```

```

step: 250, loss: 0.00011860082304337993
step: 260, loss: 0.0010028829565271735
step: 270, loss: 0.00045349393622018397
step: 280, loss: 0.00012086331844329834
step: 290, loss: 7.677663234062493e-05
step: 300, loss: 7.889031985541806e-05
step: 310, loss: 0.002103217178955674
step: 320, loss: 7.080104842316359e-05
step: 330, loss: 6.0107293393230066e-05
step: 340, loss: 5.7503493735566735e-05
step: 350, loss: 7.172356708906591e-05
step: 360, loss: 0.00011819760402431712
step: 370, loss: 0.001182957785204053
step: 380, loss: 0.05886521190404892
step: 390, loss: 6.230571307241917e-05
step: 400, loss: 0.008809964172542095
step: 410, loss: 0.00034293666249141097
step: 420, loss: 0.0004384904750622809
step: 430, loss: 0.0021308762952685356
step: 440, loss: 0.00012358918320387602
step: 450, loss: 0.0037815968971699476
step: 460, loss: 0.0002845269627869129
step: 470, loss: 0.00041539600351825356
step: 480, loss: 0.0004915300523862243
step: 490, loss: 8.974951197160408e-05
step: 500, loss: 0.0010819716844707727
step: 510, loss: 0.0003530748945195228
step: 520, loss: 0.00026843976229429245
step: 530, loss: 0.011114147491753101
/usr/local/lib/python3.7/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=0.9466852109411219, f1=0.9431870669745959,
best_f1=0.9457720588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016708907787688076
step: 10, loss: 0.01948426477611065
step: 20, loss: 0.014195389114320278
step: 30, loss: 0.00017088036111090332
step: 40, loss: 0.006438602693378925
step: 50, loss: 0.015375752933323383
step: 60, loss: 0.00011997437832178548
step: 70, loss: 0.013421637937426567
step: 80, loss: 9.747822332428768e-05
step: 90, loss: 0.006473036482930183
step: 100, loss: 0.263052374124527

```

```
step: 110, loss: 0.00010169806773774326
step: 120, loss: 0.003981973510235548
step: 130, loss: 0.000662481936160475
step: 140, loss: 0.00016246653103735298
step: 150, loss: 0.0009848610498011112
step: 160, loss: 0.0002936555538326502
step: 170, loss: 0.000219769382965751
step: 180, loss: 0.00010304337774869055
step: 190, loss: 0.00018368106975685805
step: 200, loss: 0.006168482359498739
step: 210, loss: 0.00015224417438730597
step: 220, loss: 0.00010457296593813226
step: 230, loss: 0.00036312727024778724
step: 240, loss: 0.0003135574224870652
step: 250, loss: 0.000737700960598886
step: 260, loss: 7.113766332622617e-05
step: 270, loss: 0.0006814018706791103
step: 280, loss: 0.07474581152200699
step: 290, loss: 0.0001164171117125079
step: 300, loss: 0.00033077894477173686
step: 310, loss: 0.006113688461482525
step: 320, loss: 0.0011317231692373753
step: 330, loss: 0.00031366405892185867
step: 340, loss: 9.891838271869346e-05
step: 350, loss: 8.312634599860758e-05
step: 360, loss: 0.00016560325457248837
step: 370, loss: 8.747872925596312e-05
step: 380, loss: 0.0011031798785552382
step: 390, loss: 0.003724090987816453
step: 400, loss: 5.536886601475999e-05
step: 410, loss: 0.010766529478132725
step: 420, loss: 3.809064219240099e-05
step: 430, loss: 0.0019363322062417865
step: 440, loss: 0.001766126835718751
step: 450, loss: 0.00017242228204850107
step: 460, loss: 0.0010830489918589592
step: 470, loss: 0.0005425084382295609
step: 480, loss: 0.0002078386169159785
step: 490, loss: 0.0001434189616702497
step: 500, loss: 0.000247280957410112
step: 510, loss: 0.00017916571232490242
step: 520, loss: 0.0009607716929167509
step: 530, loss: 0.0004890264826826751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9518127581459385, f1=0.9457221711131555,
best_f1=0.9457221711131555
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15281304717063904
step: 10, loss: 0.0005488135502673686
step: 20, loss: 0.00011833229655167088
step: 30, loss: 0.00012875095126219094
step: 40, loss: 0.0001311754749622196
step: 50, loss: 0.0012561503099277616
step: 60, loss: 0.00015014447853900492
step: 70, loss: 0.0014516967348754406
step: 80, loss: 0.00022052564600016922
step: 90, loss: 0.0001731603406369686
step: 100, loss: 0.0001522970269434154
step: 110, loss: 4.794376218342222e-05
step: 120, loss: 6.351996853481978e-05
step: 130, loss: 5.7657780416775495e-05
step: 140, loss: 0.0001070866928785108
step: 150, loss: 0.0011344769736751914
step: 160, loss: 0.00016292993677780032
step: 170, loss: 3.765444125747308e-05
step: 180, loss: 0.00011923282727366313
step: 190, loss: 0.00019520550267770886
step: 200, loss: 6.646794645348564e-05
step: 210, loss: 7.477445615222678e-05
step: 220, loss: 5.192526805330999e-05
step: 230, loss: 0.003367544850334525
step: 240, loss: 9.825693268794566e-05
step: 250, loss: 0.00021452023065648973
step: 260, loss: 0.022134484723210335
step: 270, loss: 0.011578948237001896
step: 280, loss: 8.06389725767076e-05
step: 290, loss: 3.143831418128684e-05
step: 300, loss: 6.337830564007163e-05
step: 310, loss: 7.052904402371496e-05
step: 320, loss: 6.582021887879819e-05
step: 330, loss: 0.00012977671576663852
step: 340, loss: 0.00014596089022234082
step: 350, loss: 3.2827108952915296e-05
step: 360, loss: 0.0879112258553505
step: 370, loss: 0.00010363533510826528
step: 380, loss: 5.940158007433638e-05
step: 390, loss: 0.00020665445481427014
step: 400, loss: 0.0001730201911414042

```

```

step: 410, loss: 0.0004893119912594557
step: 420, loss: 5.927484744461253e-05
step: 430, loss: 0.00015479352441616356
step: 440, loss: 4.6243272663559765e-05
step: 450, loss: 0.0002475048240739852
step: 460, loss: 0.0021166722290217876
step: 470, loss: 1.6848205632413737e-05
step: 480, loss: 5.9895464801229537e-05
step: 490, loss: 0.00010650509648257867
step: 500, loss: 0.0001276816619792953
step: 510, loss: 0.0001591074833413586
step: 520, loss: 0.00020353170111775398
step: 530, loss: 6.354976358124986e-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=0.9484440315838365, f1=0.945707656612529,
best_f1=0.9457221711131555
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029858911875635386
step: 10, loss: 8.451424946542829e-05
step: 20, loss: 6.18120830040425e-05
step: 30, loss: 0.0005001974059268832
step: 40, loss: 0.00045765924733132124
step: 50, loss: 0.0005454337224364281
step: 60, loss: 0.00013509212294593453
step: 70, loss: 7.321382872760296e-05
step: 80, loss: 0.00012046597839798778
step: 90, loss: 4.286606781533919e-05
step: 100, loss: 5.144132592249662e-05
step: 110, loss: 0.0006111850379966199
step: 120, loss: 6.27937915851362e-05
step: 130, loss: 0.00010485750681255013
step: 140, loss: 0.0008835049229674041
step: 150, loss: 0.00033112181699834764
step: 160, loss: 8.875672938302159e-05
step: 170, loss: 0.023241017013788223
step: 180, loss: 5.100432099425234e-05
step: 190, loss: 0.0006038694991730154
step: 200, loss: 7.722307782387361e-05
step: 210, loss: 0.0005382876843214035
step: 220, loss: 4.4883279770147055e-05
step: 230, loss: 0.008243129588663578
step: 240, loss: 0.00011697630543494597
step: 250, loss: 0.00011432875180616975
step: 260, loss: 6.324069545371458e-05

```

```

step: 270, loss: 0.00026253313990309834
step: 280, loss: 4.45301593572367e-05
step: 290, loss: 0.00013450533151626587
step: 300, loss: 7.877437019487843e-05
step: 310, loss: 3.4108095860574394e-05
step: 320, loss: 9.106918150791898e-05
step: 330, loss: 7.428489334415644e-05
step: 340, loss: 0.0004717679403256625
step: 350, loss: 6.322105036815628e-05
step: 360, loss: 7.00641394359991e-05
step: 370, loss: 0.003967850003391504
step: 380, loss: 0.00011077929229941219
step: 390, loss: 0.007773424498736858
step: 400, loss: 0.009163563139736652
step: 410, loss: 0.00014691684918943793
step: 420, loss: 0.0001253648370038718
step: 430, loss: 0.00021694770839530975
step: 440, loss: 7.650681800441816e-05
step: 450, loss: 8.312782301800326e-05
step: 460, loss: 0.003910931758582592
step: 470, loss: 7.694357918808237e-05
step: 480, loss: 5.350371066015214e-05
step: 490, loss: 7.181888213381171e-05
step: 500, loss: 0.0005302278441376984
step: 510, loss: 0.0002553272352088243
step: 520, loss: 4.613661440089345e-05
step: 530, loss: 0.001337038935162127
/usr/local/lib/python3.7/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=0.9503676470588236, f1=0.9457221711131555,
best_f1=0.9457221711131555
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.92364589869976e-05
step: 10, loss: 0.0008272346458397806
step: 20, loss: 0.008334164507687092
step: 30, loss: 0.02567734196782112
step: 40, loss: 7.372957043116912e-05
step: 50, loss: 0.00028416758868843317
step: 60, loss: 9.809016046347097e-05
step: 70, loss: 0.00017402517551090568
step: 80, loss: 7.706101314397529e-05
step: 90, loss: 4.006892777397297e-05
step: 100, loss: 2.2014344722265378e-05
step: 110, loss: 9.434700041310862e-05
step: 120, loss: 4.5960681745782495e-05

```

```

step: 130, loss: 0.0001082337403204292
step: 140, loss: 4.8634781705914065e-05
step: 150, loss: 5.625746780424379e-05
step: 160, loss: 6.87888962091878e-05
step: 170, loss: 3.403371010790579e-05
step: 180, loss: 0.00021924771135672927
step: 190, loss: 0.00040320472908206284
step: 200, loss: 0.0004814044514205307
step: 210, loss: 0.00028423621552065015
step: 220, loss: 6.287856376729906e-05
step: 230, loss: 6.557350570801646e-05
step: 240, loss: 4.4598513341043144e-05
step: 250, loss: 6.260205555008724e-05
step: 260, loss: 4.124939368921332e-05
step: 270, loss: 7.902624201960862e-05
step: 280, loss: 0.0003438879793975502
step: 290, loss: 0.00020962451526429504
step: 300, loss: 4.1178875108016655e-05
step: 310, loss: 0.03600265830755234
step: 320, loss: 6.868335913168266e-05
step: 330, loss: 6.689513975288719e-05
step: 340, loss: 8.198370778700337e-05
step: 350, loss: 0.0024500067811459303
step: 360, loss: 0.00012089385563740507
step: 370, loss: 8.330192940775305e-05
step: 380, loss: 6.710628076689318e-05
step: 390, loss: 0.003309238702058792
step: 400, loss: 0.00015595760487485677
step: 410, loss: 0.00010778185242088512
step: 420, loss: 7.987410936038941e-05
step: 430, loss: 2.5068939066841267e-05
step: 440, loss: 4.695201278082095e-05
step: 450, loss: 0.00465276837348938
step: 460, loss: 0.0011086869053542614
step: 470, loss: 4.758194336318411e-05
step: 480, loss: 0.002520486246794462
step: 490, loss: 0.006431929301470518
step: 500, loss: 0.0013435266446322203
step: 510, loss: 4.5029635657556355e-05
step: 520, loss: 0.00023590789351146668
step: 530, loss: 3.472721073194407e-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=0.950668510834486, f1=0.9474169741697417,
best_f1=0.9457221711131555
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 [00:27, 212.39it/s]

load_f1 = 0.9512867647058824

real_f1 = 0.9501845018450183

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.")
```

3072it [00:19, 159.20it/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/

```

Downloading: 100% 481/481 [00:00<00:00, 432kB/s]
 Downloading: 100% 899k/899k [00:01<00:00, 664kB/s]
 Downloading: 100% 456k/456k [00:00<00:00, 514kB/s]
 Downloading: 100% 501M/501M [00:06<00:00, 72.1MB/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: 0.48997870087623596
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4913041889667511
step: 20, loss: 0.4295283555984497
step: 30, loss: 0.4171154797077179
step: 40, loss: 0.3133029043674469
step: 50, loss: 0.4287368655204773
step: 60, loss: 0.5041627287864685
step: 70, loss: 0.3040984570980072
step: 80, loss: 0.3563336133956909
step: 90, loss: 0.23365052044391632
step: 100, loss: 0.2409820854663849
step: 110, loss: 0.26834815740585327
step: 120, loss: 0.4066017270088196
step: 130, loss: 0.22527725994586945
step: 140, loss: 0.41087082028388977
step: 150, loss: 0.21109294891357422
step: 160, loss: 0.3493611514568329
step: 170, loss: 0.10069160163402557
step: 180, loss: 0.2217995822429657
step: 190, loss: 0.6060208678245544
step: 200, loss: 0.24462489783763885
step: 210, loss: 0.34817683696746826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45701357466063347, f1=0.5109489051094891,

```

```

best_f1=0.5109489051094891
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1995459645986557
step: 10, loss: 0.04545178264379501
step: 20, loss: 0.3148389458656311
step: 30, loss: 0.28009238839149475
step: 40, loss: 0.4094841778278351
step: 50, loss: 0.2823280096054077
step: 60, loss: 0.2826738655567169
step: 70, loss: 0.1816072165966034
step: 80, loss: 0.25284072756767273
step: 90, loss: 0.10630038380622864
step: 100, loss: 0.4729175269603729
step: 110, loss: 0.2529750466346741
step: 120, loss: 0.11586282402276993
step: 130, loss: 0.2717822194099426
step: 140, loss: 0.055757392197847366
step: 150, loss: 0.3431849479675293
step: 160, loss: 0.07513298094272614
step: 170, loss: 0.30956676602363586
step: 180, loss: 0.18584804236888885
step: 190, loss: 0.3024771809577942
step: 200, loss: 0.08096962422132492
step: 210, loss: 0.15322071313858032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5911708253358925, f1=0.6490566037735849,
best_f1=0.6490566037735849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07438531517982483
step: 10, loss: 0.03942955657839775
step: 20, loss: 0.2968551516532898
step: 30, loss: 0.04832638055086136
step: 40, loss: 0.3245123028755188
step: 50, loss: 0.6663655042648315
step: 60, loss: 0.22819513082504272
step: 70, loss: 0.09533371776342392
step: 80, loss: 0.09382502734661102
step: 90, loss: 0.032218847423791885

```

```

step: 100, loss: 0.19503512978553772
step: 110, loss: 0.027850110083818436
step: 120, loss: 0.15770511329174042
step: 130, loss: 0.11313892155885696
step: 140, loss: 0.1005442664027214
step: 150, loss: 0.20582224428653717
step: 160, loss: 0.20707392692565918
step: 170, loss: 0.18214698135852814
step: 180, loss: 0.10383421927690506
step: 190, loss: 0.06752997636795044
step: 200, loss: 0.1276053637266159
step: 210, loss: 0.2559078335762024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6160337552742615, f1=0.630901287553648,
best_f1=0.630901287553648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026565004140138626
step: 10, loss: 0.1060129851102829
step: 20, loss: 0.06923678517341614
step: 30, loss: 0.13275252282619476
step: 40, loss: 0.10282894968986511
step: 50, loss: 0.11573521047830582
step: 60, loss: 0.15089291334152222
step: 70, loss: 0.05608930066227913
step: 80, loss: 0.0693850889801979
step: 90, loss: 0.04993879050016403
step: 100, loss: 0.2064078450202942
step: 110, loss: 0.2881375253200531
step: 120, loss: 0.22028443217277527
step: 130, loss: 0.1536373794078827
step: 140, loss: 0.20378366112709045
step: 150, loss: 0.12874437868595123
step: 160, loss: 0.07021689414978027
step: 170, loss: 0.03189394250512123
step: 180, loss: 0.012673833407461643
step: 190, loss: 0.13228140771389008
step: 200, loss: 0.02659166231751442
step: 210, loss: 0.3405245542526245
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6219512195121951, f1=0.6464646464646464,
best_f1=0.6464646464646464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18909864127635956
step: 10, loss: 0.022873099893331528
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.23217450082302094
step: 30, loss: 0.032679446041584015
step: 40, loss: 0.0691877007484436
step: 50, loss: 0.07784590125083923
step: 60, loss: 0.012318750843405724
step: 70, loss: 0.20921748876571655
step: 80, loss: 0.0700816810131073
step: 90, loss: 0.00959338154643774
step: 100, loss: 0.04948166385293007
step: 110, loss: 0.0314820222556591
step: 120, loss: 0.09533495455980301
step: 130, loss: 0.05852266028523445
step: 140, loss: 0.09706690907478333
step: 150, loss: 0.055425144731998444
step: 160, loss: 0.0826314389705658
step: 170, loss: 0.09565312415361404
step: 180, loss: 0.09721894562244415
step: 190, loss: 0.11385904997587204
step: 200, loss: 0.08505335450172424
step: 210, loss: 0.06761366873979568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6279999999999999, f1=0.6262230919765167,
best_f1=0.6262230919765167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0108219338580966
step: 10, loss: 0.029696546494960785
step: 20, loss: 0.013139784336090088

```

```

step: 30, loss: 0.16186951100826263
step: 40, loss: 0.02791205607354641
step: 50, loss: 0.19665439426898956
step: 60, loss: 0.1821356862783432
step: 70, loss: 0.03436017036437988
step: 80, loss: 0.13917528092861176
step: 90, loss: 0.11147172749042511
step: 100, loss: 0.15671610832214355
step: 110, loss: 0.11562304943799973
step: 120, loss: 0.021679555997252464
step: 130, loss: 0.014746226370334625
step: 140, loss: 0.04575882852077484
step: 150, loss: 0.004073515068739653
step: 160, loss: 0.024334365501999855
step: 170, loss: 0.029098201543092728
step: 180, loss: 0.04423961415886879
step: 190, loss: 0.0937507227063179
step: 200, loss: 0.22653008997440338
step: 210, loss: 0.07765081524848938
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.632183908045977, f1=0.6529080675422138,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14027981460094452
step: 10, loss: 0.01665184088051319
step: 20, loss: 0.019298618659377098
step: 30, loss: 0.04848788306117058
step: 40, loss: 0.012622109614312649
step: 50, loss: 0.0033412473276257515
step: 60, loss: 0.010022720322012901
step: 70, loss: 0.05490158870816231
step: 80, loss: 0.059461560100317
step: 90, loss: 0.02027590200304985
step: 100, loss: 0.05945947766304016
step: 110, loss: 0.08284947276115417
step: 120, loss: 0.06628581136465073
step: 130, loss: 0.08979469537734985
step: 140, loss: 0.05283847823739052
step: 150, loss: 0.035605959594249725
step: 160, loss: 0.3220904469490051

```

```

step: 170, loss: 0.06795283406972885
step: 180, loss: 0.01403819676488638
step: 190, loss: 0.05995499715209007
step: 200, loss: 0.09083256870508194
step: 210, loss: 0.2574966251850128
/usr/local/lib/python3.7/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=0.6180555555555555, f1=0.6348122866894198,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01127727422863245
step: 10, loss: 0.06511002779006958
step: 20, loss: 0.007499145809561014
step: 30, loss: 0.00513075664639473
step: 40, loss: 0.0162037443369627
step: 50, loss: 0.0009367917664349079
step: 60, loss: 0.004478253424167633
step: 70, loss: 0.06039496138691902
step: 80, loss: 0.12584325671195984
step: 90, loss: 0.018595585599541664
step: 100, loss: 0.017059920355677605
step: 110, loss: 0.04148821160197258
step: 120, loss: 0.029245909303426743
step: 130, loss: 0.012104939669370651
step: 140, loss: 0.12581849098205566
step: 150, loss: 0.15041698515415192
step: 160, loss: 0.15729652345180511
step: 170, loss: 0.06632950156927109
step: 180, loss: 0.1028166338801384
step: 190, loss: 0.022988326847553253
step: 200, loss: 0.0752691850066185
step: 210, loss: 0.1599026769399643
/usr/local/lib/python3.7/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=0.5916666666666667, f1=0.6575875486381323,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17491255700588226
step: 10, loss: 0.025704538449645042
step: 20, loss: 0.1235373392701149
step: 30, loss: 0.04950091615319252
step: 40, loss: 0.001446588896214962

```



```

step: 50, loss: 0.08267734199762344
step: 60, loss: 0.0012093264376744628
step: 70, loss: 0.01996835693717003
step: 80, loss: 0.035126298666000366
step: 90, loss: 0.036350298672914505
step: 100, loss: 0.006310855038464069
step: 110, loss: 0.03165159747004509
step: 120, loss: 0.005282060708850622
step: 130, loss: 0.08876372873783112
step: 140, loss: 0.04493141546845436
step: 150, loss: 0.015112076885998249
step: 160, loss: 0.08728092163801193
step: 170, loss: 0.08422859758138657
step: 180, loss: 0.12130557745695114
step: 190, loss: 0.0051775467582046986
step: 200, loss: 0.0027192735578864813
step: 210, loss: 0.2503019869327545
/usr/local/lib/python3.7/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=0.6290672451193059, f1=0.6425531914893616,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002190419239923358
step: 10, loss: 0.0024581067264080048
step: 20, loss: 0.006344596389681101
step: 30, loss: 0.007497986312955618
step: 40, loss: 0.005523958709090948
step: 50, loss: 0.041372478008270264
step: 60, loss: 0.004050040151923895
step: 70, loss: 0.012075434438884258
step: 80, loss: 0.011999307200312614
step: 90, loss: 0.02140599489212036
step: 100, loss: 0.13670068979263306
step: 110, loss: 0.004072379320859909
step: 120, loss: 0.017742132768034935
step: 130, loss: 0.008165913634002209
step: 140, loss: 0.007389216683804989
step: 150, loss: 0.02250286005437374
step: 160, loss: 0.01569509692490101
step: 170, loss: 0.0537242628633976
step: 180, loss: 0.002819550223648548
step: 190, loss: 0.16235309839248657
step: 200, loss: 0.03313417360186577
step: 210, loss: 0.01376986876130104
/usr/local/lib/python3.7/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=0.6117136659436009, f1=0.6570841889117043,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07864436507225037
step: 10, loss: 0.0038328575901687145
step: 20, loss: 0.05047287046909332
step: 30, loss: 0.021919287741184235
step: 40, loss: 0.060737382620573044
step: 50, loss: 0.08578158169984818
step: 60, loss: 0.0029789558611810207
step: 70, loss: 0.01155425701290369
step: 80, loss: 0.003195652738213539
step: 90, loss: 0.12952034175395966
step: 100, loss: 0.01570882275700569
step: 110, loss: 0.009769337251782417
step: 120, loss: 0.02142399549484253
step: 130, loss: 0.0017724739154800773
step: 140, loss: 0.05150997266173363
step: 150, loss: 0.01365633960813284
step: 160, loss: 0.008012956008315086
step: 170, loss: 0.003291117027401924
step: 180, loss: 0.13828854262828827
step: 190, loss: 0.011474746279418468
step: 200, loss: 0.0020464088302105665
step: 210, loss: 0.06736191362142563
/usr/local/lib/python3.7/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=0.6317907444668007, f1=0.6693069306930693,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010145921260118484
step: 10, loss: 0.0031461473554372787
step: 20, loss: 0.028789963573217392
step: 30, loss: 0.0017948143649846315
step: 40, loss: 0.0020784884691238403
step: 50, loss: 0.029871691018342972
step: 60, loss: 0.00011812973389169201
step: 70, loss: 0.0631651058793068
step: 80, loss: 0.12320003658533096
step: 90, loss: 0.0019517220789566636
step: 100, loss: 0.004767616279423237

```

```

step: 110, loss: 0.008421561680734158
step: 120, loss: 0.000731404812540859
step: 130, loss: 0.002498430898413062
step: 140, loss: 0.07283590734004974
step: 150, loss: 5.240128666628152e-05
step: 160, loss: 0.0008409464498981833
step: 170, loss: 0.1581185758113861
step: 180, loss: 0.0031294000800698996
step: 190, loss: 0.0020185718312859535
step: 200, loss: 0.027105774730443954
step: 210, loss: 0.0018127872608602047
/usr/local/lib/python3.7/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=0.6049382716049383, f1=0.6734279918864097,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001386292395181954
step: 10, loss: 0.017484283074736595
step: 20, loss: 0.0009730517631396651
step: 30, loss: 0.000765630102250725
step: 40, loss: 0.0019003867637366056
step: 50, loss: 0.1608768105506897
step: 60, loss: 0.003973386250436306
step: 70, loss: 0.008828266523778439
step: 80, loss: 0.1388845592737198
step: 90, loss: 0.06180627644062042
step: 100, loss: 0.002189646940678358
step: 110, loss: 0.010883791372179985
step: 120, loss: 0.012864246964454651
step: 130, loss: 0.004337153863161802
step: 140, loss: 0.0472613163292408
step: 150, loss: 0.0014988286420702934
step: 160, loss: 0.0032776256557554007
step: 170, loss: 0.0009046490304172039
step: 180, loss: 0.0006068386719562113
step: 190, loss: 0.01724270172417164
step: 200, loss: 0.0022915585432201624
step: 210, loss: 0.009688161313533783
/usr/local/lib/python3.7/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=0.6119733924611973, f1=0.6535087719298245,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001220264588482678
step: 10, loss: 0.009963509626686573
step: 20, loss: 0.002201601630076766
step: 30, loss: 0.006598723586648703
step: 40, loss: 0.010194428265094757
step: 50, loss: 0.0009412798681296408
step: 60, loss: 0.042046111077070236
step: 70, loss: 0.0007149601588025689
step: 80, loss: 0.009382717311382294
step: 90, loss: 0.0002094357623718679
step: 100, loss: 0.0004947561537846923
step: 110, loss: 0.02598848007619381
step: 120, loss: 0.000522151414770633
step: 130, loss: 0.008197974413633347
step: 140, loss: 0.0007587178843095899
step: 150, loss: 0.0013836967991665006
step: 160, loss: 0.0003197974874638021
step: 170, loss: 0.022127648815512657
step: 180, loss: 0.010767209343612194
step: 190, loss: 0.002963251667097211
step: 200, loss: 0.0012701739324256778
step: 210, loss: 0.0018642987124621868
/usr/local/lib/python3.7/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=0.5815602836879433, f1=0.6374133949191686,
best_f1=0.6529080675422138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008687230874784291
step: 10, loss: 0.0007351858075708151
step: 20, loss: 0.014517168514430523
step: 30, loss: 0.006976981181651354
step: 40, loss: 0.0044340514577925205
step: 50, loss: 0.0023712425027042627
step: 60, loss: 0.01247647125273943
step: 70, loss: 0.0026460441295057535
step: 80, loss: 0.009348362684249878
step: 90, loss: 0.003322921460494399
step: 100, loss: 0.003566832048818469
step: 110, loss: 0.003746827132999897
step: 120, loss: 0.0021776396315544844
step: 130, loss: 0.0006331360200420022
step: 140, loss: 0.0037612980231642723
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.005297404248267412

```

```

step: 160, loss: 0.04192742332816124
step: 170, loss: 0.007231942377984524
step: 180, loss: 0.0010233931243419647
step: 190, loss: 0.0012944458285346627
step: 200, loss: 0.0018351764883846045
step: 210, loss: 0.0026944917626678944
/usr/local/lib/python3.7/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=0.5977528089887639, f1=0.6403508771929826,
best_f1=0.6529080675422138
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 [00:10, 223.89it/s]
load_f1 = 0.6234309623430961
real_f1 = 0.6029723991507431
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 142.40it/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: 0.45684614777565
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4493238031864166
step: 20, loss: 0.2752213478088379
step: 30, loss: 0.41212794184684753
step: 40, loss: 0.2367752492427826
step: 50, loss: 0.3174777626991272
step: 60, loss: 0.46438562870025635
step: 70, loss: 0.4132888913154602
step: 80, loss: 0.1885468065738678
step: 90, loss: 0.366184264421463
step: 100, loss: 0.44298481941223145
step: 110, loss: 0.2425331175327301
step: 120, loss: 0.3414851725101471
step: 130, loss: 0.31588447093963623
step: 140, loss: 0.16788288950920105
step: 150, loss: 0.3125526010990143
step: 160, loss: 0.23601576685905457
step: 170, loss: 0.36983436346054077
step: 180, loss: 0.18523991107940674
step: 190, loss: 0.1793089509010315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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: 0.4122224748134613
step: 10, loss: 0.336031436920166
step: 20, loss: 0.6138032078742981
step: 30, loss: 0.25099024176597595
step: 40, loss: 0.6064199805259705
step: 50, loss: 0.31144025921821594
step: 60, loss: 0.46132487058639526
step: 70, loss: 0.32125362753868103
step: 80, loss: 0.16638898849487305
step: 90, loss: 0.3202342987060547
step: 100, loss: 0.27338308095932007
step: 110, loss: 0.37246009707450867
step: 120, loss: 0.23750533163547516
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5357649326324463
step: 140, loss: 0.31338179111480713
step: 150, loss: 0.298373818397522
step: 160, loss: 0.30970993638038635
step: 170, loss: 0.23507127165794373
step: 180, loss: 0.17659997940063477
step: 190, loss: 0.2488642781972885
/usr/local/lib/python3.7/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=0.17216770740410348, 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: 0.38108476996421814
step: 10, loss: 0.3692318797111511
step: 20, loss: 0.4792785942554474
step: 30, loss: 0.31184256076812744
step: 40, loss: 0.08475299924612045
step: 50, loss: 0.3816800117492676
step: 60, loss: 0.16129866242408752
step: 70, loss: 0.38006138801574707
step: 80, loss: 0.31251874566078186
step: 90, loss: 0.37946730852127075
step: 100, loss: 0.5558768510818481
```



```

step: 110, loss: 0.6842050552368164
step: 120, loss: 0.3851817846298218
step: 130, loss: 0.16854235529899597
step: 140, loss: 0.3782051205635071
step: 150, loss: 0.3191101551055908
step: 160, loss: 0.6105090975761414
step: 170, loss: 0.44137823581695557
step: 180, loss: 0.3887787461280823
step: 190, loss: 0.1634047031402588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.24206779897212982
step: 10, loss: 0.23090168833732605
step: 20, loss: 0.3039539158344269
step: 30, loss: 0.24358269572257996
step: 40, loss: 0.5479764938354492
step: 50, loss: 0.23073604702949524
step: 60, loss: 0.4028238356113434
step: 70, loss: 0.30401793122291565
step: 80, loss: 0.2559671998023987
step: 90, loss: 0.18734267354011536
step: 100, loss: 0.30260276794433594
step: 110, loss: 0.38322123885154724
step: 120, loss: 0.24145051836967468
step: 130, loss: 0.4542091190814972
step: 140, loss: 0.3117624819278717
step: 150, loss: 0.23353931307792664
step: 160, loss: 0.3145466446876526
step: 170, loss: 0.44398269057273865
step: 180, loss: 0.3715381324291229
step: 190, loss: 0.16088518500328064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.3939831256866455
step: 10, loss: 0.4041374623775482
step: 20, loss: 0.1746644228696823

```

```

step: 30, loss: 0.11528606712818146
step: 40, loss: 0.3044593632221222
step: 50, loss: 0.5036059021949768
step: 60, loss: 0.24943989515304565
step: 70, loss: 0.37247180938720703
step: 80, loss: 0.3766288161277771
step: 90, loss: 0.3040508031845093
step: 100, loss: 0.4515043795108795
step: 110, loss: 0.40569445490837097
step: 120, loss: 0.2382752001285553
step: 130, loss: 0.5534902811050415
step: 140, loss: 0.38410645723342896
step: 150, loss: 0.3119066655635834
step: 160, loss: 0.16470129787921906
step: 170, loss: 0.3876091241836548
step: 180, loss: 0.24886056780815125
step: 190, loss: 0.3114710748195648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.303220272064209
step: 10, loss: 0.24216626584529877
step: 20, loss: 0.30625173449516296
step: 30, loss: 0.47788986563682556
step: 40, loss: 0.24016405642032623
step: 50, loss: 0.3078833818435669
step: 60, loss: 0.4517662227153778
step: 70, loss: 0.3009762763977051
step: 80, loss: 0.3138546645641327
step: 90, loss: 0.24650812149047852
step: 100, loss: 0.4730309844017029
step: 110, loss: 0.23017454147338867
step: 120, loss: 0.4669490456581116
step: 130, loss: 0.5234544277191162
step: 140, loss: 0.18550655245780945
step: 150, loss: 0.37263572216033936
step: 160, loss: 0.3796660006046295
step: 170, loss: 0.39084935188293457
step: 180, loss: 0.1794779747724533
step: 190, loss: 0.2968323528766632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

epoch 6: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.3165417015552521
step: 10, loss: 0.3188442885875702
step: 20, loss: 0.3117900490760803
step: 30, loss: 0.18920032680034637
step: 40, loss: 0.3205529749393463
step: 50, loss: 0.09418133646249771
step: 60, loss: 0.1583709865808487
step: 70, loss: 0.15615281462669373
step: 80, loss: 0.23891586065292358
step: 90, loss: 0.23813077807426453
step: 100, loss: 0.5815250277519226
step: 110, loss: 0.44547414779663086
step: 120, loss: 0.3908972144126892
step: 130, loss: 0.3119947910308838
step: 140, loss: 0.23819875717163086
step: 150, loss: 0.31248435378074646
step: 160, loss: 0.37979093194007874
step: 170, loss: 0.3137996196746826
step: 180, loss: 0.23692891001701355
step: 190, loss: 0.30611366033554077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.23810426890850067
step: 10, loss: 0.3788575828075409
step: 20, loss: 0.23317654430866241
step: 30, loss: 0.3787134885787964
step: 40, loss: 0.16957710683345795
step: 50, loss: 0.3240986466407776
step: 60, loss: 0.44645819067955017
step: 70, loss: 0.23801563680171967
step: 80, loss: 0.38123172521591187
step: 90, loss: 0.23900309205055237
step: 100, loss: 0.310198575258255
step: 110, loss: 0.3848428428173065
step: 120, loss: 0.3797273337841034
step: 130, loss: 0.3063049614429474
step: 140, loss: 0.3975204825401306

```

```

step: 150, loss: 0.32162395119667053
step: 160, loss: 0.189456969499588
step: 170, loss: 0.23549029231071472
step: 180, loss: 0.3161049783229828
step: 190, loss: 0.3131697475910187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.2483399212360382
step: 10, loss: 0.10454124212265015
step: 20, loss: 0.31123417615890503
step: 30, loss: 0.17999160289764404
step: 40, loss: 0.30457767844200134
step: 50, loss: 0.3774728775024414
step: 60, loss: 0.38415947556495667
step: 70, loss: 0.15339705348014832
step: 80, loss: 0.39186543226242065
step: 90, loss: 0.6830716729164124
step: 100, loss: 0.3751152455806732
step: 110, loss: 0.3876441717147827
step: 120, loss: 0.5906596779823303
step: 130, loss: 0.3193076550960541
step: 140, loss: 0.31318211555480957
step: 150, loss: 0.2402643859386444
step: 160, loss: 0.23108047246932983
step: 170, loss: 0.4656487703323364
step: 180, loss: 0.31344932317733765
step: 190, loss: 0.24173028767108917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.24157670140266418
step: 10, loss: 0.16649997234344482
step: 20, loss: 0.31188341975212097
step: 30, loss: 0.3777919113636017
step: 40, loss: 0.3151613473892212
step: 50, loss: 0.0954471305012703
step: 60, loss: 0.31191641092300415

```

```

step: 70, loss: 0.3156558871269226
step: 80, loss: 0.39581236243247986
step: 90, loss: 0.1628040224313736
step: 100, loss: 0.23311279714107513
step: 110, loss: 0.23152361810207367
step: 120, loss: 0.38951870799064636
step: 130, loss: 0.5116056203842163
step: 140, loss: 0.3700537383556366
step: 150, loss: 0.31404438614845276
step: 160, loss: 0.24050620198249817
step: 170, loss: 0.3792445659637451
step: 180, loss: 0.23547926545143127
step: 190, loss: 0.16329023241996765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.5439802408218384
step: 10, loss: 0.09766943752765656
step: 20, loss: 0.234677255153656
step: 30, loss: 0.23904335498809814
step: 40, loss: 0.08698387444019318
step: 50, loss: 0.3179932236671448
step: 60, loss: 0.23470327258110046
step: 70, loss: 0.6170253753662109
step: 80, loss: 0.3769751489162445
step: 90, loss: 0.313738077878952
step: 100, loss: 0.1752840131521225
step: 110, loss: 0.25041577219963074
step: 120, loss: 0.3166857659816742
step: 130, loss: 0.611445963382721
step: 140, loss: 0.45719197392463684
step: 150, loss: 0.3157779276371002
step: 160, loss: 0.3160971999168396
step: 170, loss: 0.4431585669517517
step: 180, loss: 0.3163556754589081
step: 190, loss: 0.10849975794553757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.37746915221214294
step: 10, loss: 0.30805283784866333
step: 20, loss: 0.5388367772102356
step: 30, loss: 0.3825536370277405
step: 40, loss: 0.3061036765575409
step: 50, loss: 0.5294281244277954
step: 60, loss: 0.4576943516731262
step: 70, loss: 0.3884565234184265
step: 80, loss: 0.3777627944946289
step: 90, loss: 0.45182347297668457
step: 100, loss: 0.3179713487625122
step: 110, loss: 0.4598005414009094
step: 120, loss: 0.16987088322639465
step: 130, loss: 0.24756833910942078
step: 140, loss: 0.3845612704753876
step: 150, loss: 0.3776514232158661
step: 160, loss: 0.243598073720932
step: 170, loss: 0.455814391374588
step: 180, loss: 0.31439805030822754
step: 190, loss: 0.16956643760204315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.31558936834335327
step: 10, loss: 0.3124021291732788
step: 20, loss: 0.40269720554351807
step: 30, loss: 0.3921460509300232
step: 40, loss: 0.24705329537391663
step: 50, loss: 0.24713429808616638
step: 60, loss: 0.31046929955482483
step: 70, loss: 0.1652788519859314
step: 80, loss: 0.31080567836761475
step: 90, loss: 0.3138391673564911
step: 100, loss: 0.16162168979644775
step: 110, loss: 0.6043486595153809
step: 120, loss: 0.16696317493915558
step: 130, loss: 0.23469585180282593
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.3128146529197693
step: 150, loss: 0.37754741311073303
step: 160, loss: 0.4501582980155945
step: 170, loss: 0.10869179666042328

```

```

step: 180, loss: 0.24573028087615967
step: 190, loss: 0.3109624683856964
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.31468477845191956
step: 10, loss: 0.2418137639760971
step: 20, loss: 0.3882666528224945
step: 30, loss: 0.10351122170686722
step: 40, loss: 0.3116414546966553
step: 50, loss: 0.3831918239593506
step: 60, loss: 0.31706956028938293
step: 70, loss: 0.244558185338974
step: 80, loss: 0.37839147448539734
step: 90, loss: 0.17216727137565613
step: 100, loss: 0.3111822307109833
step: 110, loss: 0.4454711079597473
step: 120, loss: 0.3064422011375427
step: 130, loss: 0.3144829571247101
step: 140, loss: 0.17180778086185455
step: 150, loss: 0.4518009424209595
step: 160, loss: 0.38834747672080994
step: 170, loss: 0.2445935755968094
step: 180, loss: 0.31021249294281006
step: 190, loss: 0.5273017287254333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.5428845882415771
step: 10, loss: 0.23364295065402985
step: 20, loss: 0.2421085685491562
step: 30, loss: 0.3135065734386444
step: 40, loss: 0.6660134792327881
step: 50, loss: 0.3101087212562561
step: 60, loss: 0.176253080368042
step: 70, loss: 0.2364198863506317
step: 80, loss: 0.2434941679239273
step: 90, loss: 0.3148539662361145

```

```

step: 100, loss: 0.39001893997192383
step: 110, loss: 0.4545298218727112
step: 120, loss: 0.24000994861125946
step: 130, loss: 0.31167182326316833
step: 140, loss: 0.445420503616333
step: 150, loss: 0.1669899970293045
step: 160, loss: 0.30531638860702515
step: 170, loss: 0.7443451285362244
step: 180, loss: 0.31143617630004883
step: 190, loss: 0.45634958148002625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:13, 150.10it/s]
load_f1 = 0.17216770740410348
real_f1 = 0.17216770740410348

```



```
733it [00:00, 3545.39it/s]/usr/local/lib/python3.7/dist-  
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
4398it [00:30, 141.95it/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 \n  
  --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 \n  
  --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: 0.5224134922027588
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.43570470809936523
step: 20, loss: 0.3386564254760742
step: 30, loss: 0.4864845871925354
step: 40, loss: 0.5513207912445068
step: 50, loss: 0.3058820068836212
step: 60, loss: 0.6161268949508667
step: 70, loss: 0.31029242277145386
step: 80, loss: 0.22977179288864136
step: 90, loss: 0.19007262587547302
step: 100, loss: 0.18375064432621002
step: 110, loss: 0.4068055748939514
step: 120, loss: 0.332864910364151
step: 130, loss: 0.33719414472579956
step: 140, loss: 0.3461642861366272
step: 150, loss: 0.3085441589355469

```

```

step: 160, loss: 0.38738834857940674
step: 170, loss: 0.3276815116405487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.32663947343826294
step: 10, loss: 0.4489154517650604
step: 20, loss: 0.311973512172699
step: 30, loss: 0.3248671889305115
step: 40, loss: 0.09652520716190338
step: 50, loss: 0.4534939229488373
step: 60, loss: 0.19815103709697723
step: 70, loss: 0.4873529076576233
step: 80, loss: 0.2504025995731354
step: 90, loss: 0.23449283838272095
step: 100, loss: 0.5136857628822327
step: 110, loss: 0.26196008920669556
step: 120, loss: 0.2552545666694641
step: 130, loss: 0.5388697385787964
step: 140, loss: 0.5181161761283875
step: 150, loss: 0.43537476658821106
step: 160, loss: 0.4303972125053406
step: 170, loss: 0.37198373675346375
/usr/local/lib/python3.7/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: 0.6010770201683044
step: 10, loss: 0.30089306831359863
step: 20, loss: 0.23111656308174133
step: 30, loss: 0.23997394740581512
step: 40, loss: 0.3856006860733032
step: 50, loss: 0.5737312436103821
step: 60, loss: 0.31186139583587646
step: 70, loss: 0.23026776313781738

```

```

step: 80, loss: 0.3747130036354065
step: 90, loss: 0.5237874984741211
step: 100, loss: 0.30894196033477783
step: 110, loss: 0.16972661018371582
step: 120, loss: 0.5781854391098022
step: 130, loss: 0.512714147567749
step: 140, loss: 0.43094122409820557
step: 150, loss: 0.19514276087284088
step: 160, loss: 0.17397497594356537
step: 170, loss: 0.3072023391723633
/usr/local/lib/python3.7/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: 0.4033510982990265
step: 10, loss: 0.5287095308303833
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.23550793528556824
step: 30, loss: 0.4498196244239807
step: 40, loss: 0.24383828043937683
step: 50, loss: 0.38628238439559937
step: 60, loss: 0.6455332636833191
step: 70, loss: 0.32093605399131775
step: 80, loss: 0.46076318621635437
step: 90, loss: 0.32136183977127075
step: 100, loss: 0.3771967589855194
step: 110, loss: 0.4422902464866638
step: 120, loss: 0.469236820936203
step: 130, loss: 0.3115960359573364
step: 140, loss: 0.23853453993797302
step: 150, loss: 0.6925938725471497
step: 160, loss: 0.13616178929805756
step: 170, loss: 0.23973330855369568
/usr/local/lib/python3.7/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: 0.38206979632377625
step: 10, loss: 0.3002461791038513
step: 20, loss: 0.3522549271583557

```

```

step: 30, loss: 0.2979086637496948
step: 40, loss: 0.23861420154571533
step: 50, loss: 0.23809799551963806
step: 60, loss: 0.3088415563106537
step: 70, loss: 0.32274994254112244
step: 80, loss: 0.07553055882453918
step: 90, loss: 0.6234117746353149
step: 100, loss: 0.25978484749794006
step: 110, loss: 0.26905491948127747
step: 120, loss: 0.13239283859729767
step: 130, loss: 0.23931871354579926
step: 140, loss: 0.16747640073299408
step: 150, loss: 0.2987862527370453
step: 160, loss: 0.2363199144601822
step: 170, loss: 0.37887316942214966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2315227070347284, f1=0.23260247855100097,
best_f1=0.23260247855100097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3040986657142639
step: 10, loss: 0.10812019556760788
step: 20, loss: 0.35027897357940674
step: 30, loss: 0.25723958015441895
step: 40, loss: 0.450551837682724
step: 50, loss: 0.23476389050483704
step: 60, loss: 0.3711256682872772
step: 70, loss: 0.3677956759929657
step: 80, loss: 0.18618735671043396
step: 90, loss: 0.3080659508705139
step: 100, loss: 0.1614590436220169
step: 110, loss: 0.457817405462265
step: 120, loss: 0.4307559132575989
step: 130, loss: 0.5421786308288574
step: 140, loss: 0.25127190351486206
step: 150, loss: 0.14907178282737732
step: 160, loss: 0.3210066556930542
step: 170, loss: 0.22419333457946777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.28079470198675494, f1=0.26720647773279355,
best_f1=0.26720647773279355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26566728949546814
step: 10, loss: 0.4422398805618286
step: 20, loss: 0.46231937408447266
step: 30, loss: 0.4653851389884949
step: 40, loss: 0.12200221419334412
step: 50, loss: 0.3856482207775116
step: 60, loss: 0.53426194190979
step: 70, loss: 0.33202752470970154
step: 80, loss: 0.17256517708301544
step: 90, loss: 0.4491376578807831
step: 100, loss: 0.27592507004737854
step: 110, loss: 0.3402920365333557
step: 120, loss: 0.3485150635242462
step: 130, loss: 0.24777241051197052
step: 140, loss: 0.1400597095489502
step: 150, loss: 0.27838245034217834
step: 160, loss: 0.35033002495765686
step: 170, loss: 0.3041612505912781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.31558935361216733, f1=0.27184466019417475,
best_f1=0.27184466019417475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4057004451751709
step: 10, loss: 0.45555421710014343
step: 20, loss: 0.2114461213350296
step: 30, loss: 0.2379002571105957
step: 40, loss: 0.20314805209636688
step: 50, loss: 0.3260626792907715
step: 60, loss: 0.3439945876598358
step: 70, loss: 0.13149790465831757
step: 80, loss: 0.23917439579963684

```

```

step: 90, loss: 0.3075391352176666
step: 100, loss: 0.16885268688201904
step: 110, loss: 0.3430001139640808
step: 120, loss: 0.45840683579444885
step: 130, loss: 0.22471940517425537
step: 140, loss: 0.36341404914855957
step: 150, loss: 0.21012230217456818
step: 160, loss: 0.04184854030609131
step: 170, loss: 0.32737305760383606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5111111111111111, f1=0.5107913669064748,
best_f1=0.5107913669064748
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15873926877975464
step: 10, loss: 0.33107584714889526
step: 20, loss: 0.2887733578681946
step: 30, loss: 0.12931105494499207
step: 40, loss: 0.2327919900417328
step: 50, loss: 0.12048786878585815
step: 60, loss: 0.2560717463493347
step: 70, loss: 0.27690640091896057
step: 80, loss: 0.0652107521891594
step: 90, loss: 0.3683861792087555
step: 100, loss: 0.22373196482658386
step: 110, loss: 0.1421109288930893
step: 120, loss: 0.4050123691558838
step: 130, loss: 0.19916735589504242
step: 140, loss: 0.11499611288309097
step: 150, loss: 0.3012973964214325
step: 160, loss: 0.1798066943883896
step: 170, loss: 0.183517724275589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5905172413793105, f1=0.6004140786749482,
best_f1=0.6004140786749482

```

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

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

```
step: 0, loss: 0.14966821670532227  
step: 10, loss: 0.11000782251358032  
step: 20, loss: 0.11786771565675735  
step: 30, loss: 0.33142703771591187  
step: 40, loss: 0.3250548839569092  
step: 50, loss: 0.31806644797325134  
step: 60, loss: 0.3754658102989197  
step: 70, loss: 0.14593444764614105  
step: 80, loss: 0.22078245878219604  
step: 90, loss: 0.07587716728448868  
step: 100, loss: 0.28691521286964417  
step: 110, loss: 0.10530348122119904  
step: 120, loss: 0.13329382240772247  
step: 130, loss: 0.02025538496673107  
step: 140, loss: 0.056089192628860474  
step: 150, loss: 0.2473679482936859  
step: 160, loss: 0.0777050256729126  
step: 170, loss: 0.25564202666282654
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not 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 10: dev_f1=0.6150234741784038, f1=0.6605080831408776,  
best_f1=0.6605080831408776
```

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

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

```
step: 0, loss: 0.2711876630783081  
step: 10, loss: 0.20467163622379303  
step: 20, loss: 0.11962679028511047  
step: 30, loss: 0.3253336250782013  
step: 40, loss: 0.052735086530447006  
step: 50, loss: 0.1387292593717575  
step: 60, loss: 0.36775025725364685  
step: 70, loss: 0.17952892184257507  
step: 80, loss: 0.07757637649774551  
step: 90, loss: 0.09068766236305237  
step: 100, loss: 0.3823051154613495  
step: 110, loss: 0.14194858074188232  
step: 120, loss: 0.21721746027469635  
step: 130, loss: 0.16307522356510162  
step: 140, loss: 0.30255889892578125
```



```

step: 150, loss: 0.043906837701797485
step: 160, loss: 0.07221173495054245
step: 170, loss: 0.1796836405992508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6372549019607844, f1=0.65083135391924,
best_f1=0.65083135391924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23494018614292145
step: 10, loss: 0.1253843605518341
step: 20, loss: 0.04874030500650406
step: 30, loss: 0.23375681042671204
step: 40, loss: 0.09184136986732483
step: 50, loss: 0.09575656801462173
step: 60, loss: 0.11857355386018753
step: 70, loss: 0.02821836806833744
step: 80, loss: 0.06997161358594894
step: 90, loss: 0.1720402091741562
step: 100, loss: 0.03144475072622299
step: 110, loss: 0.2017461061477661
step: 120, loss: 0.09840720891952515
step: 130, loss: 0.19714058935642242
step: 140, loss: 0.15901316702365875
step: 150, loss: 0.05177291855216026
step: 160, loss: 0.3843386173248291
step: 170, loss: 0.34623900055885315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6453089244851259, f1=0.6802721088435375,
best_f1=0.6802721088435375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053392939269542694
step: 10, loss: 0.06714648753404617
step: 20, loss: 0.05164986103773117

```

```

step: 30, loss: 0.1639413982629776
step: 40, loss: 0.15190184116363525
step: 50, loss: 0.08667688816785812
step: 60, loss: 0.05180596560239792
step: 70, loss: 0.2251470983028412
step: 80, loss: 0.048560578376054764
step: 90, loss: 0.0936276987195015
step: 100, loss: 0.04376468434929848
step: 110, loss: 0.19664238393306732
step: 120, loss: 0.03118598647415638
step: 130, loss: 0.023097224533557892
step: 140, loss: 0.24996815621852875
step: 150, loss: 0.1854085624217987
step: 160, loss: 0.14240983128547668
step: 170, loss: 0.061763741075992584
/usr/local/lib/python3.7/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=0.6368159203980098, f1=0.7014218009478673,
best_f1=0.6802721088435375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.060375623404979706
step: 10, loss: 0.04751350358128548
step: 20, loss: 0.19994410872459412
step: 30, loss: 0.15146008133888245
step: 40, loss: 0.06977199018001556
step: 50, loss: 0.03141534700989723
step: 60, loss: 0.16314558684825897
step: 70, loss: 0.07681863009929657
step: 80, loss: 0.02540522627532482
step: 90, loss: 0.09598355740308762
step: 100, loss: 0.5308424234390259
step: 110, loss: 0.18308450281620026
step: 120, loss: 0.15432685613632202
step: 130, loss: 0.137806236743927
step: 140, loss: 0.14560078084468842
step: 150, loss: 0.055990926921367645
step: 160, loss: 0.09214131534099579
step: 170, loss: 0.2506602108478546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6499999999999999, f1=0.6821052631578948,
best_f1=0.6821052631578948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02796945534646511
step: 10, loss: 0.06627184897661209
step: 20, loss: 0.017916670069098473
step: 30, loss: 0.15545184910297394
step: 40, loss: 0.04530680552124977
step: 50, loss: 0.05775542929768562
step: 60, loss: 0.040231604129076004
step: 70, loss: 0.18604779243469238
step: 80, loss: 0.038952842354774475
step: 90, loss: 0.1946214735507965
step: 100, loss: 0.19153544306755066
step: 110, loss: 0.1723310798406601
step: 120, loss: 0.04669990763068199
step: 130, loss: 0.2823021411895752
step: 140, loss: 0.08788474649190903
step: 150, loss: 0.023760942742228508
step: 160, loss: 0.10568135231733322
step: 170, loss: 0.06069212406873703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6711409395973154, f1=0.7022222222222223,
best_f1=0.7022222222222223
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, 227.92it/s]
load_f1 = 0.5632458233890215
real_f1 = 0.558252427184466
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 [00:30, 142.21it/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: 0.5979142189025879
```

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: 0.43974819779396057
```

```
step: 20, loss: 0.5192098617553711
```

```
step: 30, loss: 0.3724230229854584
```

```
step: 40, loss: 0.32762035727500916
```

```
step: 50, loss: 0.5776663422584534
```

```
step: 60, loss: 0.47590023279190063
```

```

step: 70, loss: 0.44028374552726746
step: 80, loss: 0.45457297563552856
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.24901996552944183
step: 100, loss: 0.14870940148830414
step: 110, loss: 0.36679503321647644
step: 120, loss: 0.2251165509223938
step: 130, loss: 0.13468994200229645
step: 140, loss: 0.015590266324579716
step: 150, loss: 0.14457473158836365
step: 160, loss: 0.05450360104441643
step: 170, loss: 0.18422673642635345
step: 180, loss: 0.06268929690122604
step: 190, loss: 0.07920637726783752
step: 200, loss: 0.05996327102184296
step: 210, loss: 0.06184128299355507
step: 220, loss: 0.10505848377943039
step: 230, loss: 0.0589505136013031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9288793103448276, f1=0.9393605292171995,
best_f1=0.9393605292171995
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010152212344110012
step: 10, loss: 0.10773037374019623
step: 20, loss: 0.02154688537120819
step: 30, loss: 0.04101619869470596
step: 40, loss: 0.04598267376422882
step: 50, loss: 0.020777912810444832
step: 60, loss: 0.10956314951181412
step: 70, loss: 0.033505573868751526
step: 80, loss: 0.04454544186592102
step: 90, loss: 0.02890956401824951
step: 100, loss: 0.029684094712138176
step: 110, loss: 0.03516850993037224
step: 120, loss: 0.038363974541425705
step: 130, loss: 0.06068016588687897
step: 140, loss: 0.011324673891067505
step: 150, loss: 0.2201681286096573
step: 160, loss: 0.009802674874663353
step: 170, loss: 0.001951185055077076

```

```

step: 180, loss: 0.015689648687839508
step: 190, loss: 0.03518122062087059
step: 200, loss: 0.21628141403198242
step: 210, loss: 0.029797684401273727
step: 220, loss: 0.003201527986675501
step: 230, loss: 0.0010501743527129292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9610678531701891, f1=0.958659217877095,
best_f1=0.958659217877095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05671127885580063
step: 10, loss: 0.06249891594052315
step: 20, loss: 0.00806399155408144
step: 30, loss: 0.004731875378638506
step: 40, loss: 0.01776645891368389
step: 50, loss: 0.03138623759150505
step: 60, loss: 0.0053806910291314125
step: 70, loss: 0.004765905439853668
step: 80, loss: 0.13458870351314545
step: 90, loss: 0.006155447568744421
step: 100, loss: 0.020609281957149506
step: 110, loss: 0.01170342043042183
step: 120, loss: 0.07286889851093292
step: 130, loss: 0.011447442695498466
step: 140, loss: 0.014716355130076408
step: 150, loss: 0.03486248478293419
step: 160, loss: 0.029231654480099678
step: 170, loss: 0.002287516836076975
step: 180, loss: 0.01033734530210495
step: 190, loss: 0.05286332592368126
step: 200, loss: 0.006833489052951336
step: 210, loss: 0.004358714446425438
step: 220, loss: 0.06377564370632172
step: 230, loss: 0.0032149769831448793
/usr/local/lib/python3.7/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=0.9597315436241611, f1=0.9515219842164601,
best_f1=0.958659217877095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007411854807287455
step: 10, loss: 0.005643106065690517
step: 20, loss: 0.0024781033862382174
step: 30, loss: 0.028024787083268166
step: 40, loss: 0.03441194072365761
step: 50, loss: 0.00836469978094101
step: 60, loss: 0.015653930604457855
step: 70, loss: 0.0014090933836996555
step: 80, loss: 0.0026050955057144165
step: 90, loss: 0.02394113875925541
step: 100, loss: 0.025860587134957314
step: 110, loss: 0.005856078118085861
step: 120, loss: 0.02509853057563305
step: 130, loss: 0.01922258734703064
step: 140, loss: 0.001737719401717186
step: 150, loss: 0.0020551877096295357
step: 160, loss: 0.0025387529749423265
step: 170, loss: 0.002203996991738677
step: 180, loss: 0.15785184502601624
step: 190, loss: 0.0018684528768062592
step: 200, loss: 0.005214964505285025
step: 210, loss: 0.0036381399258971214
step: 220, loss: 0.004178878851234913
step: 230, loss: 0.0012273527681827545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9707865168539327, f1=0.9579067121729239,
best_f1=0.9579067121729239
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007847091765142977
step: 10, loss: 0.002923512365669012
step: 20, loss: 0.004008636809885502
step: 30, loss: 0.003671307349577546
step: 40, loss: 0.01601197011768818
step: 50, loss: 0.013792771846055984
step: 60, loss: 0.13889168202877045
step: 70, loss: 0.001808722736313939
step: 80, loss: 0.006557678338140249
step: 90, loss: 0.022779548540711403

```



```

step: 100, loss: 0.0005174692487344146
step: 110, loss: 0.0012832535430788994
step: 120, loss: 0.010965442284941673
step: 130, loss: 0.0007203927962109447
step: 140, loss: 0.016500404104590416
step: 150, loss: 0.0036914353258907795
step: 160, loss: 0.00745379738509655
step: 170, loss: 0.15137679874897003
step: 180, loss: 0.08554988354444504
step: 190, loss: 0.03533066064119339
step: 200, loss: 0.0019591364543884993
step: 210, loss: 0.00474750017747283
step: 220, loss: 0.002946150256320834
step: 230, loss: 0.004235965199768543
/usr/local/lib/python3.7/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=0.9648924122310306, f1=0.9576174112256587,
best_f1=0.9579067121729239
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006558767636306584
step: 10, loss: 0.004351991694420576
step: 20, loss: 0.0018532747635617852
step: 30, loss: 0.0010179925011470914
step: 40, loss: 0.000429611187428236
step: 50, loss: 0.0004918023478239775
step: 60, loss: 0.0014585691969841719
step: 70, loss: 0.04166656732559204
step: 80, loss: 0.005797484889626503
step: 90, loss: 0.0028149655554443598
step: 100, loss: 0.0005226911744102836
step: 110, loss: 0.0058654919266700745
step: 120, loss: 0.00041720259469002485
step: 130, loss: 0.000572561810258776
step: 140, loss: 0.00837730523198843
step: 150, loss: 0.0012084355112165213
step: 160, loss: 0.009739595465362072
step: 170, loss: 0.069234699010849
step: 180, loss: 0.013407486490905285
step: 190, loss: 0.004736336413770914
step: 200, loss: 0.004168704617768526
step: 210, loss: 0.0013167030410841107
step: 220, loss: 0.010824727825820446
step: 230, loss: 0.0023995577357709408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9720044792833147, f1=0.9704545454545453,
best_f1=0.9704545454545453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002695647068321705
step: 10, loss: 0.0009634889429435134
step: 20, loss: 0.00034907704684883356
step: 30, loss: 0.0004072723095305264
step: 40, loss: 0.0014796770410612226
step: 50, loss: 0.0012225176906213164
step: 60, loss: 0.0012191931018605828
step: 70, loss: 0.0007110853912308812
step: 80, loss: 0.0014118656981736422
step: 90, loss: 0.0006033413228578866
step: 100, loss: 0.0005564825260080397
step: 110, loss: 0.0004649000184144825
step: 120, loss: 0.0009599559125490487
step: 130, loss: 0.0015707287238910794
step: 140, loss: 0.0002783865202218294
step: 150, loss: 0.010302890092134476
step: 160, loss: 0.00032064612605609
step: 170, loss: 0.009474161081016064
step: 180, loss: 0.0006306513678282499
step: 190, loss: 0.015309489332139492
step: 200, loss: 0.003561084158718586
step: 210, loss: 0.0010394882410764694
step: 220, loss: 0.0005286672385409474
step: 230, loss: 0.000827137497253716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9774266365688488, 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: 0.0006643295637331903
step: 10, loss: 0.0021995059214532375

```

```

step: 20, loss: 0.003238845383748412
step: 30, loss: 0.001720731146633625
step: 40, loss: 0.0032200815621763468
step: 50, loss: 0.0005496556404978037
step: 60, loss: 0.0007529343129135668
step: 70, loss: 0.0004768540384247899
step: 80, loss: 0.022370925173163414
step: 90, loss: 0.0021529102232307196
step: 100, loss: 0.001088427845388651
step: 110, loss: 0.004459598567336798
step: 120, loss: 0.00044783749035559595
step: 130, loss: 0.0009647779515944421
step: 140, loss: 0.0005263531347736716
step: 150, loss: 0.058689579367637634
step: 160, loss: 0.000521199603099376
step: 170, loss: 0.013489334844052792
step: 180, loss: 0.0009617999894544482
step: 190, loss: 0.022518504410982132
step: 200, loss: 0.05441886559128761
step: 210, loss: 0.002001391025260091
step: 220, loss: 0.0008731010602787137
step: 230, loss: 0.00037444400368258357
/usr/local/lib/python3.7/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=0.9774266365688488, f1=0.9726027397260274,
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: 0.0010390623938292265
step: 10, loss: 0.001784015097655356
step: 20, loss: 0.0009925502818077803
step: 30, loss: 0.002172334585338831
step: 40, loss: 0.0011534662917256355
step: 50, loss: 0.0009222966036759317
step: 60, loss: 0.0009638680494390428
step: 70, loss: 0.056951526552438736
step: 80, loss: 0.0024483134038746357
step: 90, loss: 0.02031221054494381
step: 100, loss: 0.0005038008093833923
step: 110, loss: 0.0005953013896942139
step: 120, loss: 0.010871898382902145
step: 130, loss: 0.0012828364269807935
step: 140, loss: 0.000825157854706049
step: 150, loss: 0.004624711349606514
step: 160, loss: 0.00707043893635273
step: 170, loss: 0.0002836265484802425

```

```

step: 180, loss: 0.010008223354816437
step: 190, loss: 0.00020636072440538555
step: 200, loss: 0.0002189340302720666
step: 210, loss: 0.0013835112331435084
step: 220, loss: 0.004271875601261854
step: 230, loss: 0.0005290705012157559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9786276715410572, f1=0.9785310734463276,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008619112195447087
step: 10, loss: 0.003274794900789857
step: 20, loss: 0.0008746837265789509
step: 30, loss: 0.0004080715880263597
step: 40, loss: 0.0032385983504354954
step: 50, loss: 0.000601286708842963
step: 60, loss: 0.0012090844102203846
step: 70, loss: 0.008310348726809025
step: 80, loss: 0.001742631895467639
step: 90, loss: 0.0005156337283551693
step: 100, loss: 0.0003988728567492217
step: 110, loss: 0.0009523688349872828
step: 120, loss: 0.0005309226689860225
step: 130, loss: 0.0008586474577896297
step: 140, loss: 0.00034501292975619435
step: 150, loss: 0.0011109693441540003
step: 160, loss: 0.00014913732593413442
step: 170, loss: 0.0003010663203895092
step: 180, loss: 0.0010187654988840222
step: 190, loss: 0.00040814929525367916
step: 200, loss: 0.0010485327802598476
step: 210, loss: 0.05217184126377106
step: 220, loss: 0.0009399367263540626
step: 230, loss: 0.00031116779427975416
/usr/local/lib/python3.7/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=0.9761634506242906, f1=0.9727272727272728,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002826851268764585
step: 10, loss: 0.00042476426460780203
step: 20, loss: 0.00031422049505636096
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0005101457936689258
step: 40, loss: 5.4267125960905105e-05
step: 50, loss: 0.00017756257147993892
step: 60, loss: 0.014154382981359959
step: 70, loss: 0.0007547697750851512
step: 80, loss: 0.025964342057704926
step: 90, loss: 0.18743599951267242
step: 100, loss: 0.002803550101816654
step: 110, loss: 0.0003954597341362387
step: 120, loss: 0.0030291101429611444
step: 130, loss: 0.0004425695224199444
step: 140, loss: 0.0018375799991190434
step: 150, loss: 0.0009974241256713867
step: 160, loss: 0.0027335137128829956
step: 170, loss: 0.0016421866603195667
step: 180, loss: 0.0013575946213677526
step: 190, loss: 0.0012692271266132593
step: 200, loss: 0.006110710557550192
step: 210, loss: 0.0008249871898442507
step: 220, loss: 0.0012552484404295683
step: 230, loss: 0.000638590136077255
/usr/local/lib/python3.7/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=0.9763779527559054, f1=0.9684684684684683,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010599942179396749
step: 10, loss: 0.0003862686571665108
step: 20, loss: 0.0010855947621166706
step: 30, loss: 0.0470668226480484
step: 40, loss: 0.0011672339169308543
step: 50, loss: 0.0007003373466432095
step: 60, loss: 0.001618965994566679
step: 70, loss: 0.00033749762224033475
step: 80, loss: 0.0005412072059698403
step: 90, loss: 0.00038790807593613863
step: 100, loss: 0.0005790450377389789
step: 110, loss: 0.00028899015160277486
step: 120, loss: 0.0031947060488164425

```

```

step: 130, loss: 0.0017420303774997592
step: 140, loss: 0.0006504691555164754
step: 150, loss: 0.0006733782356604934
step: 160, loss: 0.004730419255793095
step: 170, loss: 0.0002682568447198719
step: 180, loss: 0.0008831224404275417
step: 190, loss: 0.001867205137386918
step: 200, loss: 0.001542786369100213
step: 210, loss: 0.0003145380178466439
step: 220, loss: 0.0007877483149059117
step: 230, loss: 0.0014309256803244352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819004524886877, f1=0.97272727272728,
best_f1=0.97272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007371920510195196
step: 10, loss: 0.0008396076736971736
step: 20, loss: 0.0010678916005417705
step: 30, loss: 0.009816392324864864
step: 40, loss: 0.0007472258294001222
step: 50, loss: 0.0018767430447041988
step: 60, loss: 0.00037059857277199626
step: 70, loss: 0.000156951486133039
step: 80, loss: 9.270204463973641e-05
step: 90, loss: 0.0001970810699276626
step: 100, loss: 0.0007784431800246239
step: 110, loss: 0.000536116014700383
step: 120, loss: 0.00039518577978014946
step: 130, loss: 0.0004875545564573258
step: 140, loss: 0.001444666413590312
step: 150, loss: 6.913291872479022e-05
step: 160, loss: 0.0023040231317281723
step: 170, loss: 0.0006595898885279894
step: 180, loss: 0.029028838500380516
step: 190, loss: 0.00024532689712941647
step: 200, loss: 4.773797627422027e-05
step: 210, loss: 0.0006029165233485401
step: 220, loss: 0.00015535883721895516
step: 230, loss: 0.0001300879375776276
/usr/local/lib/python3.7/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=0.9806157354618015, f1=0.9702517162471395,
best_f1=0.9727272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011333864676998928
step: 10, loss: 0.00010812225809786469
step: 20, loss: 0.00012684284592978656
step: 30, loss: 0.000278992869425565
step: 40, loss: 0.00039110853685997427
step: 50, loss: 9.495155973127112e-05
step: 60, loss: 0.0003506613720674068
step: 70, loss: 0.00046431980445049703
step: 80, loss: 0.00010987724090227857
step: 90, loss: 0.0008656770805828273
step: 100, loss: 0.0003941546310670674
step: 110, loss: 0.0006772708729840815
step: 120, loss: 3.7277703086147085e-05
step: 130, loss: 0.0006340260151773691
step: 140, loss: 0.0005039240932092071
step: 150, loss: 6.623627268709242e-05
step: 160, loss: 0.0001622192357899621
step: 170, loss: 0.00020950843463651836
step: 180, loss: 0.00021053769160062075
step: 190, loss: 0.0011629292275756598
step: 200, loss: 0.00019406831415835768
step: 210, loss: 0.00043045455822721124
step: 220, loss: 0.00043261994142085314
step: 230, loss: 0.00013276255049277097
/usr/local/lib/python3.7/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=0.9786276715410572, f1=0.9774266365688488,
best_f1=0.9727272727272728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0215901006013155
step: 10, loss: 0.00041371717816218734
step: 20, loss: 0.0003284417907707393
step: 30, loss: 0.00021591248514596373
step: 40, loss: 9.542836050968617e-05
step: 50, loss: 0.0013763055903837085
step: 60, loss: 0.008747619576752186
step: 70, loss: 9.226417023455724e-05
step: 80, loss: 0.00013256481906864792

```

```

step: 90, loss: 7.146182178985327e-05
step: 100, loss: 6.883819150971249e-05
step: 110, loss: 0.0001811133261071518
step: 120, loss: 0.023283224552869797
step: 130, loss: 7.077717600623146e-05
step: 140, loss: 0.00029884057585150003
step: 150, loss: 0.00014342715439852327
step: 160, loss: 0.011614852584898472
step: 170, loss: 5.613160828943364e-05
step: 180, loss: 0.0003933763946406543
step: 190, loss: 0.00021164222562219948
step: 200, loss: 9.433337254449725e-05
step: 210, loss: 0.05304417759180069
step: 220, loss: 0.0009045694023370743
step: 230, loss: 0.000119248048576992
/usr/local/lib/python3.7/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=0.9785794813979707, f1=0.9762174405436014,
best_f1=0.9727272727272728
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:17, 143.28it/s]
load_f1 = 0.9831271091113611
real_f1 = 0.9820224719101124
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 [00:33, 130.31it/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 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: 0.6222481727600098
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.4687756597995758
step: 20, loss: 0.25222116708755493
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.398689866065979
step: 40, loss: 0.2763591706752777
step: 50, loss: 0.42517539858818054
step: 60, loss: 0.23702993988990784
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.16185471415519714
step: 80, loss: 0.13969337940216064
step: 90, loss: 0.14175757765769958
step: 100, loss: 0.21059833467006683
step: 110, loss: 0.20550650358200073
step: 120, loss: 0.20189253985881805
step: 130, loss: 0.09982144832611084
step: 140, loss: 0.3466428518295288
step: 150, loss: 0.21334826946258545
step: 160, loss: 0.2653432786464691
step: 170, loss: 0.06638039648532867
step: 180, loss: 0.10087010264396667
step: 190, loss: 0.12132685631513596
step: 200, loss: 0.05892464518547058

```

```

step: 210, loss: 0.11885136365890503
step: 220, loss: 0.03264876455068588
step: 230, loss: 0.14015783369541168
step: 240, loss: 0.029243450611829758
step: 250, loss: 0.13082334399223328
step: 260, loss: 0.25772568583488464
step: 270, loss: 0.22044460475444794
step: 280, loss: 0.08009412884712219
step: 290, loss: 0.11576090008020401
step: 300, loss: 0.0910399928689003
step: 310, loss: 0.10463692992925644
step: 320, loss: 0.04567507654428482
step: 330, loss: 0.03423400595784187
step: 340, loss: 0.32875901460647583
step: 350, loss: 0.10525234788656235
step: 360, loss: 0.034358568489551544
step: 370, loss: 0.04010371118783951
step: 380, loss: 0.09423958510160446
step: 390, loss: 0.013202971778810024
step: 400, loss: 0.03537483140826225
step: 410, loss: 0.11058027297258377
step: 420, loss: 0.03165515139698982
step: 430, loss: 0.03906974941492081
step: 440, loss: 0.20854276418685913
step: 450, loss: 0.07616932690143585
step: 460, loss: 0.05421217903494835
step: 470, loss: 0.04189828410744667
step: 480, loss: 0.14764080941677094
step: 490, loss: 0.1152690127491951
step: 500, loss: 0.07191697508096695
step: 510, loss: 0.11227802187204361
step: 520, loss: 0.09530756622552872
step: 530, loss: 0.007885810919106007
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9314685314685315, f1=0.9260628465804066,
best_f1=0.9260628465804066
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1345495879650116
step: 10, loss: 0.06485123187303543
step: 20, loss: 0.022376805543899536

```

step: 30, loss: 0.019989600405097008
step: 40, loss: 0.030318697914481163
step: 50, loss: 0.05083683878183365
step: 60, loss: 0.010921316221356392
step: 70, loss: 0.016290972009301186
step: 80, loss: 0.06810081005096436
step: 90, loss: 0.023998968303203583
step: 100, loss: 0.17576728761196136
step: 110, loss: 0.031835075467824936
step: 120, loss: 0.029723048210144043
step: 130, loss: 0.0018795274663716555
step: 140, loss: 0.015264054760336876
step: 150, loss: 0.07465682923793793
step: 160, loss: 0.04463962838053703
step: 170, loss: 0.035091616213321686
step: 180, loss: 0.012480552308261395
step: 190, loss: 0.019205499440431595
step: 200, loss: 0.20750173926353455
step: 210, loss: 0.04304223507642746
step: 220, loss: 0.0034228952135890722
step: 230, loss: 0.03806084766983986
step: 240, loss: 0.013469639234244823
step: 250, loss: 0.0539759062230587
step: 260, loss: 0.009317677468061447
step: 270, loss: 0.05423774570226669
step: 280, loss: 0.0889991968870163
step: 290, loss: 0.01610414683818817
step: 300, loss: 0.04930848628282547
step: 310, loss: 0.034769393503665924
step: 320, loss: 0.039739955216646194
step: 330, loss: 0.0157471876591444
step: 340, loss: 0.3224872052669525
step: 350, loss: 0.0028161462396383286
step: 360, loss: 0.07619332522153854
step: 370, loss: 0.020131580531597137
step: 380, loss: 0.12902800738811493
step: 390, loss: 0.017865633592009544
step: 400, loss: 0.07320193946361542
step: 410, loss: 0.04803461208939552
step: 420, loss: 0.09432290494441986
step: 430, loss: 0.12849515676498413
step: 440, loss: 0.016255797818303108
step: 450, loss: 0.07359765470027924
step: 460, loss: 0.034134168177843094
step: 470, loss: 0.02348903939127922
step: 480, loss: 0.019897034391760826
step: 490, loss: 0.06281629204750061
step: 500, loss: 0.005218866281211376

```

step: 510, loss: 0.04502669349312782
step: 520, loss: 0.3545747399330139
step: 530, loss: 0.11252402514219284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.945387792565397, f1=0.940045766590389,
best_f1=0.940045766590389
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13286632299423218
step: 10, loss: 0.021036136895418167
step: 20, loss: 0.005271764937788248
step: 30, loss: 0.046679720282554626
step: 40, loss: 0.02264568768441677
step: 50, loss: 0.004098015837371349
step: 60, loss: 0.007866010069847107
step: 70, loss: 0.0028933004941791296
step: 80, loss: 0.09491503238677979
step: 90, loss: 0.011383970268070698
step: 100, loss: 0.017856242135167122
step: 110, loss: 0.0744975283741951
step: 120, loss: 0.10477468371391296
step: 130, loss: 0.07511209696531296
step: 140, loss: 0.005986340809613466
step: 150, loss: 0.025110263377428055
step: 160, loss: 0.011649766005575657
step: 170, loss: 0.03445901721715927
step: 180, loss: 0.06447825580835342
step: 190, loss: 0.0288309995085001
step: 200, loss: 0.03712981939315796
step: 210, loss: 0.006330684293061495
step: 220, loss: 0.14269395172595978
step: 230, loss: 0.03933484107255936
step: 240, loss: 0.05436641350388527
step: 250, loss: 0.051975067704916
step: 260, loss: 0.009552464820444584
step: 270, loss: 0.001095696585252881
step: 280, loss: 0.07883300632238388
step: 290, loss: 0.01242312602698803
step: 300, loss: 0.08091383427381516
step: 310, loss: 0.11877026408910751
step: 320, loss: 0.027559446170926094

```

```

step: 330, loss: 0.010814286768436432
step: 340, loss: 0.007905038073658943
step: 350, loss: 0.07953839004039764
step: 360, loss: 0.013201374560594559
step: 370, loss: 0.08290015161037445
step: 380, loss: 0.0625743642449379
step: 390, loss: 0.00530386483296752
step: 400, loss: 0.1795840561389923
step: 410, loss: 0.11029128730297089
step: 420, loss: 0.009505809284746647
step: 430, loss: 0.034392740577459335
step: 440, loss: 0.12730289995670319
step: 450, loss: 0.041595183312892914
step: 460, loss: 0.08564400672912598
step: 470, loss: 0.07376062870025635
step: 480, loss: 0.08614934980869293
step: 490, loss: 0.03138057887554169
step: 500, loss: 0.027612466365098953
step: 510, loss: 0.020161088556051254
step: 520, loss: 0.11835591495037079
step: 530, loss: 0.0006343710119836032
/usr/local/lib/python3.7/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=0.9171058936272161, f1=0.9103117505995204,
best_f1=0.940045766590389
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03359047695994377
step: 10, loss: 0.02946394868195057
step: 20, loss: 0.06286229938268661
step: 30, loss: 0.021447593346238136
step: 40, loss: 0.011991795152425766
step: 50, loss: 0.040909405797719955
step: 60, loss: 0.0020738327875733376
step: 70, loss: 0.006451255176216364
step: 80, loss: 0.19707857072353363
step: 90, loss: 0.11332274973392487
step: 100, loss: 0.0032006357796490192
step: 110, loss: 0.07925282418727875
step: 120, loss: 0.006335715297609568
step: 130, loss: 0.02408979833126068
step: 140, loss: 0.009111779741942883
step: 150, loss: 0.0224230345338583
step: 160, loss: 0.06643597781658173
step: 170, loss: 0.01785655878484249
step: 180, loss: 0.04790192097425461

```

```

step: 190, loss: 0.013291759416460991
step: 200, loss: 0.09718431532382965
step: 210, loss: 0.0013410215033218265
step: 220, loss: 0.0045060706324875355
step: 230, loss: 0.004652773030102253
step: 240, loss: 0.021170655265450478
step: 250, loss: 0.09899384528398514
step: 260, loss: 0.01007918082177639
step: 270, loss: 0.14631766080856323
step: 280, loss: 0.0077437409199774265
step: 290, loss: 0.043924491852521896
step: 300, loss: 0.003729134565219283
step: 310, loss: 0.005497942678630352
step: 320, loss: 0.028839675709605217
step: 330, loss: 0.011832996271550655
step: 340, loss: 0.025748729705810547
step: 350, loss: 0.08173467218875885
step: 360, loss: 0.07610231637954712
step: 370, loss: 0.0010103541426360607
step: 380, loss: 0.0018285701517015696
step: 390, loss: 0.0003827608306892216
step: 400, loss: 0.01530643180012703
step: 410, loss: 0.05396891385316849
step: 420, loss: 0.012416386976838112
step: 430, loss: 0.05873722955584526
step: 440, loss: 0.027591612190008163
step: 450, loss: 0.02614790014922619
step: 460, loss: 0.01663120836019516
step: 470, loss: 0.000201158516574651
step: 480, loss: 0.028284600004553795
step: 490, loss: 0.014871789142489433
step: 500, loss: 0.0753554180264473
step: 510, loss: 0.16413117945194244
step: 520, loss: 0.018794331699609756
step: 530, loss: 0.1182691901922226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9493258949325895, f1=0.9389846297158826,
best_f1=0.9389846297158826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012742032995447516

```

step: 10, loss: 0.002428963780403137
step: 20, loss: 0.0007411343394778669
step: 30, loss: 0.03116689808666706
step: 40, loss: 0.0008420202066190541
step: 50, loss: 0.01961652934551239
step: 60, loss: 0.01903395727276802
step: 70, loss: 0.007721927482634783
step: 80, loss: 0.00021950635709799826
step: 90, loss: 0.04792984575033188
step: 100, loss: 0.0018149006646126509
step: 110, loss: 0.05102889984846115
step: 120, loss: 0.02522006258368492
step: 130, loss: 0.008450550027191639
step: 140, loss: 0.015193958766758442
step: 150, loss: 0.0010988936992362142
step: 160, loss: 0.011728701181709766
step: 170, loss: 0.07097039371728897
step: 180, loss: 0.010517891496419907
step: 190, loss: 0.004783803131431341
step: 200, loss: 0.0015025314642116427
step: 210, loss: 0.002335903001949191
step: 220, loss: 0.0003187531838193536
step: 230, loss: 0.0016635481733828783
step: 240, loss: 0.003451660042628646
step: 250, loss: 0.2541583776473999
step: 260, loss: 0.004620358347892761
step: 270, loss: 0.0017597691621631384
step: 280, loss: 0.0026466932613402605
step: 290, loss: 0.010382218286395073
step: 300, loss: 0.006498419679701328
step: 310, loss: 0.007942970842123032
step: 320, loss: 0.007986573502421379
step: 330, loss: 0.027338039129972458
step: 340, loss: 0.00043381229625083506
step: 350, loss: 0.00021743483375757933
step: 360, loss: 0.00011790457210736349
step: 370, loss: 0.0005832333117723465
step: 380, loss: 0.00010575290070846677
step: 390, loss: 0.000689418229740113
step: 400, loss: 0.00046610485878773034
step: 410, loss: 0.21178539097309113
step: 420, loss: 0.05504967272281647
step: 430, loss: 0.020293308421969414
step: 440, loss: 0.00025272637140005827
step: 450, loss: 0.022501662373542786
step: 460, loss: 0.016251619905233383
step: 470, loss: 0.0300800409168005
step: 480, loss: 0.04442698135972023


```

step: 490, loss: 0.008555860258638859
step: 500, loss: 0.020220598205924034
step: 510, loss: 0.0030984412878751755
step: 520, loss: 0.09443891048431396
step: 530, loss: 0.04454914852976799
/usr/local/lib/python3.7/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=0.946135831381733, f1=0.9366852886405959,
best_f1=0.9389846297158826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012378288432955742
step: 10, loss: 0.00033582249307073653
step: 20, loss: 0.00044621567940339446
step: 30, loss: 0.0006944040069356561
step: 40, loss: 0.00028168881544843316
step: 50, loss: 0.00013020010374020785
step: 60, loss: 0.002224575961008668
step: 70, loss: 0.0006728843436576426
step: 80, loss: 0.00030954796238802373
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.00302052590996027
step: 100, loss: 0.004413262475281954
step: 110, loss: 0.010607964359223843
step: 120, loss: 0.004564314614981413
step: 130, loss: 0.008211871609091759
step: 140, loss: 0.0005124518647789955
step: 150, loss: 0.00041549309389665723
step: 160, loss: 0.0005557386903092265
step: 170, loss: 0.0006152383284643292
step: 180, loss: 0.00030789183801971376
step: 190, loss: 0.018026508390903473
step: 200, loss: 0.013017880730330944
step: 210, loss: 0.002529399236664176
step: 220, loss: 0.0009187771938741207
step: 230, loss: 0.010343356058001518
step: 240, loss: 0.0036010381300002337
step: 250, loss: 0.0199010968208313
step: 260, loss: 0.00034361970028840005
step: 270, loss: 0.0024585069622844458
step: 280, loss: 0.01404501497745514
step: 290, loss: 0.0012409951305016875
step: 300, loss: 0.0019567168783396482
step: 310, loss: 0.03300861269235611
step: 320, loss: 0.08929108828306198
step: 330, loss: 0.0006775101064704359

```

```

step: 340, loss: 0.0005133011727593839
step: 350, loss: 0.001238225377164781
step: 360, loss: 0.12678059935569763
step: 370, loss: 0.0024996446445584297
step: 380, loss: 0.005308326333761215
step: 390, loss: 0.005457165651023388
step: 400, loss: 0.0033769418951123953
step: 410, loss: 0.00023514588247053325
step: 420, loss: 0.000825725554022938
step: 430, loss: 0.04784563183784485
step: 440, loss: 0.019069159403443336
step: 450, loss: 0.0169255081564188
step: 460, loss: 0.0011115088127553463
step: 470, loss: 0.0013937588082626462
step: 480, loss: 0.033165231347084045
step: 490, loss: 0.011094223707914352
step: 500, loss: 0.009856962598860264
step: 510, loss: 0.02275203913450241
step: 520, loss: 0.0018366053700447083
step: 530, loss: 0.0054027908481657505
/usr/local/lib/python3.7/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=0.9489322191272052, f1=0.9282385834109972,
best_f1=0.9389846297158826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02364286221563816
step: 10, loss: 0.0012537947623059154
step: 20, loss: 0.013326931744813919
step: 30, loss: 0.005746346898376942
step: 40, loss: 0.003037018235772848
step: 50, loss: 0.0034999907948076725
step: 60, loss: 0.013255074620246887
step: 70, loss: 0.003875237889587879
step: 80, loss: 0.00021355500211939216
step: 90, loss: 0.0007093104650266469
step: 100, loss: 0.00033450796036049724
step: 110, loss: 0.00016909802798181772
step: 120, loss: 0.0001925992692122236
step: 130, loss: 6.409981870092452e-05
step: 140, loss: 0.00012576766312122345
step: 150, loss: 0.0003964120987802744
step: 160, loss: 0.0001914207241497934
step: 170, loss: 0.02457437478005886
step: 180, loss: 0.05306071788072586
step: 190, loss: 0.0078406548127532

```

```

step: 200, loss: 0.0013947099214419723
step: 210, loss: 0.0010821755276992917
step: 220, loss: 7.665294833714142e-05
step: 230, loss: 5.325497841113247e-05
step: 240, loss: 0.0007383369375020266
step: 250, loss: 0.0009319951059296727
step: 260, loss: 0.0007277869153767824
step: 270, loss: 0.00045324431266635656
step: 280, loss: 0.0028129240963608027
step: 290, loss: 0.019961830228567123
step: 300, loss: 0.000522296701092273
step: 310, loss: 0.0015573303680866957
step: 320, loss: 0.004006346222013235
step: 330, loss: 0.0001732376404106617
step: 340, loss: 0.0011409217258915305
step: 350, loss: 0.000482320407172665
step: 360, loss: 0.0032405827660113573
step: 370, loss: 0.000371068948879838
step: 380, loss: 0.0003760383988264948
step: 390, loss: 0.0003449520154390484
step: 400, loss: 0.01152227632701397
step: 410, loss: 0.017998071387410164
step: 420, loss: 0.010868764482438564
step: 430, loss: 0.00969739817082882
step: 440, loss: 0.008055963553488255
step: 450, loss: 0.0021054937969893217
step: 460, loss: 0.00025737719261087477
step: 470, loss: 0.12727707624435425
step: 480, loss: 0.04502425342798233
step: 490, loss: 0.005454534664750099
step: 500, loss: 0.005036966875195503
step: 510, loss: 0.0007516186451539397
step: 520, loss: 0.00130177428945899
step: 530, loss: 0.0004153829941060394
/usr/local/lib/python3.7/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=0.9480580252690688, f1=0.9345099860659545,
best_f1=0.9389846297158826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000719383591786027
step: 10, loss: 0.001111445832066238
step: 20, loss: 0.0005145479226484895
step: 30, loss: 0.00022879788593854755
step: 40, loss: 0.0004138169751968235
step: 50, loss: 0.00011804657697211951

```

step: 60, loss: 0.00014096482482273132
step: 70, loss: 0.00020961584232281893
step: 80, loss: 0.0007571913301944733
step: 90, loss: 0.0004878333129454404
step: 100, loss: 0.00010170792666031048
step: 110, loss: 0.0002680437173694372
step: 120, loss: 0.00017596494581084698
step: 130, loss: 0.000594042707234621
step: 140, loss: 0.028394442051649094
step: 150, loss: 9.882164158625528e-05
step: 160, loss: 0.0005771723226644099
step: 170, loss: 0.05886177346110344
step: 180, loss: 0.0006831549108028412
step: 190, loss: 0.05451106280088425
step: 200, loss: 0.0024644858203828335
step: 210, loss: 0.02610117569565773
step: 220, loss: 0.0005867881118319929
step: 230, loss: 0.05286649987101555
step: 240, loss: 0.0004326918860897422
step: 250, loss: 0.001568696927279234
step: 260, loss: 4.635624645743519e-05
step: 270, loss: 0.0008891703910194337
step: 280, loss: 5.683256677002646e-05
step: 290, loss: 0.00039420576649717987
step: 300, loss: 4.868861651630141e-05
step: 310, loss: 0.01667185313999653
step: 320, loss: 0.015372291207313538
step: 330, loss: 8.668813825352117e-05
step: 340, loss: 0.030918069183826447
step: 350, loss: 0.0005122103029862046
step: 360, loss: 0.02135869860649109
step: 370, loss: 0.00046935916179791093
step: 380, loss: 7.199988613137975e-05
step: 390, loss: 0.005224170163273811
step: 400, loss: 0.012328331358730793
step: 410, loss: 0.00037162890657782555
step: 420, loss: 0.0003336295485496521
step: 430, loss: 0.004995232447981834
step: 440, loss: 0.0006729665910825133
step: 450, loss: 0.0013991784071549773
step: 460, loss: 0.0007602116675116122
step: 470, loss: 0.0890517458319664
step: 480, loss: 0.001935733132995665
step: 490, loss: 0.001832183450460434
step: 500, loss: 0.0021512589883059263
step: 510, loss: 0.0004155548813287169
step: 520, loss: 0.0002134311362169683
step: 530, loss: 0.06730292737483978

```
/usr/local/lib/python3.7/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=0.946135831381733, f1=0.9313543599257884,
best_f1=0.9389846297158826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.99349225719925e-05
step: 10, loss: 0.003384467214345932
step: 20, loss: 0.0003640954091679305
step: 30, loss: 0.022027384489774704
step: 40, loss: 0.0007071195286698639
step: 50, loss: 0.0004872626159340143
step: 60, loss: 0.044036976993083954
step: 70, loss: 0.00039884159923531115
step: 80, loss: 0.0035217730328440666
step: 90, loss: 0.058980464935302734
step: 100, loss: 0.00024146059877239168
step: 110, loss: 0.017645955085754395
step: 120, loss: 0.01122195366770029
step: 130, loss: 0.00042916418169625103
step: 140, loss: 0.04524214193224907
step: 150, loss: 0.0024953153915703297
step: 160, loss: 0.002722739474847913
step: 170, loss: 0.004447183571755886
step: 180, loss: 0.0005699030007235706
step: 190, loss: 0.0008797499467618763
step: 200, loss: 0.0002817089552991092
step: 210, loss: 0.0004353294207248837
step: 220, loss: 0.0002471976913511753
step: 230, loss: 0.0003008613712154329
step: 240, loss: 0.00014502019621431828
step: 250, loss: 0.01972167380154133
step: 260, loss: 8.071066986303777e-05
step: 270, loss: 0.0007145727286115289
step: 280, loss: 0.001318154507316649
step: 290, loss: 0.00256825122050941
step: 300, loss: 0.009756739251315594
step: 310, loss: 0.07482026517391205
step: 320, loss: 0.001100457040593028
step: 330, loss: 0.0008170292130671442
step: 340, loss: 0.030764590948820114
step: 350, loss: 0.03370002284646034
step: 360, loss: 0.0011907437583431602
step: 370, loss: 0.0006864645984023809
step: 380, loss: 0.0008852713508531451
step: 390, loss: 0.00019388024520594627
```

```

step: 400, loss: 0.07121815532445908
step: 410, loss: 0.0004907197435386479
step: 420, loss: 0.0021747625432908535
step: 430, loss: 0.0029067066498100758
step: 440, loss: 0.0034474933054298162
step: 450, loss: 0.001085328753106296
step: 460, loss: 0.00248893559910357
step: 470, loss: 0.00030727911507710814
step: 480, loss: 0.00023368133406620473
step: 490, loss: 0.002837395528331399
step: 500, loss: 0.0020042634569108486
step: 510, loss: 0.07942042499780655
step: 520, loss: 0.0050955661572515965
step: 530, loss: 0.004820825532078743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9507139567019807, f1=0.9393382352941178,
best_f1=0.9393382352941178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002078496851027012
step: 10, loss: 0.0016287385951727629
step: 20, loss: 0.0028432002291083336
step: 30, loss: 0.0013165846467018127
step: 40, loss: 0.0002627279900480062
step: 50, loss: 0.00015552231343463063
step: 60, loss: 0.00027377717196941376
step: 70, loss: 0.00010161295358557254
step: 80, loss: 0.00016607315046712756
step: 90, loss: 0.00027528207283467054
step: 100, loss: 0.0004273467347957194
step: 110, loss: 0.06555613875389099
step: 120, loss: 0.00032432383159175515
step: 130, loss: 5.75586236664094e-05
step: 140, loss: 9.942797623807564e-05
step: 150, loss: 8.366393740288913e-05
step: 160, loss: 0.010257037356495857
step: 170, loss: 0.00015636299212928861
step: 180, loss: 0.007361717522144318
step: 190, loss: 0.0002685925574041903
step: 200, loss: 0.026290684938430786
step: 210, loss: 0.0005974387167952955

```

```

step: 220, loss: 0.0004261165449861437
step: 230, loss: 0.00011873942275997251
step: 240, loss: 0.026001915335655212
step: 250, loss: 0.00013411729014478624
step: 260, loss: 0.0005076061352156103
step: 270, loss: 0.00016831133689265698
step: 280, loss: 0.0004130890010856092
step: 290, loss: 9.534809942124411e-05
step: 300, loss: 0.0005956512759439647
step: 310, loss: 0.04515087604522705
step: 320, loss: 0.0005667465156875551
step: 330, loss: 0.00016088529082480818
step: 340, loss: 0.00012835106463171542
step: 350, loss: 0.01476207934319973
step: 360, loss: 7.130565063562244e-05
step: 370, loss: 0.00011892869952134788
step: 380, loss: 0.0013288382906466722
step: 390, loss: 7.815597200533375e-05
step: 400, loss: 8.470557077089325e-05
step: 410, loss: 7.426032971125096e-05
step: 420, loss: 7.495099271181971e-05
step: 430, loss: 0.00012859527487307787
step: 440, loss: 6.671746814390644e-05
step: 450, loss: 8.037292718654498e-05
step: 460, loss: 7.380392344202846e-05
step: 470, loss: 0.00348391174338758
step: 480, loss: 0.00032199989072978497
step: 490, loss: 0.00113195541780442
step: 500, loss: 0.00016940620844252408
step: 510, loss: 8.528976468369365e-05
step: 520, loss: 0.0020911958999931812
step: 530, loss: 0.02310180477797985
/usr/local/lib/python3.7/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=0.9478584729981377, f1=0.9385474860335196,
best_f1=0.9393382352941178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.427281863987446e-05
step: 10, loss: 0.00019317913393024355
step: 20, loss: 0.0002714560541789979
step: 30, loss: 8.975001401267946e-05
step: 40, loss: 4.819026435143314e-05
step: 50, loss: 7.279706187546253e-05
step: 60, loss: 0.002001057378947735
step: 70, loss: 0.0006370561895892024

```

step: 80, loss: 0.0013881586492061615
step: 90, loss: 0.0013227590825408697
step: 100, loss: 4.989479930372909e-05
step: 110, loss: 0.00024061260046437383
step: 120, loss: 0.00026660694857127964
step: 130, loss: 0.00014726493100170046
step: 140, loss: 0.00013828124792780727
step: 150, loss: 0.0002049892646027729
step: 160, loss: 0.002390161156654358
step: 170, loss: 0.00021986589126754552
step: 180, loss: 0.0036795083433389664
step: 190, loss: 0.0006850940990261734
step: 200, loss: 0.0001277425471926108
step: 210, loss: 0.00020645701442845166
step: 220, loss: 0.0019502323120832443
step: 230, loss: 5.174801481189206e-05
step: 240, loss: 0.0021937438286840916
step: 250, loss: 0.00014462783292401582
step: 260, loss: 0.003253370989114046
step: 270, loss: 0.007317772600799799
step: 280, loss: 0.0002235369465779513
step: 290, loss: 0.00018187062232755125
step: 300, loss: 0.00013526109978556633
step: 310, loss: 0.019326163455843925
step: 320, loss: 0.0006501174648292363
step: 330, loss: 7.270820788107812e-05
step: 340, loss: 0.031086456030607224
step: 350, loss: 0.00013255765952635556
step: 360, loss: 0.003987668082118034
step: 370, loss: 0.0007082860101945698
step: 380, loss: 0.00010662278509698808
step: 390, loss: 0.0008805729448795319
step: 400, loss: 9.22640465432778e-05
step: 410, loss: 0.0012554700952023268
step: 420, loss: 0.0010459921322762966
step: 430, loss: 0.0014280856121331453
step: 440, loss: 0.00014079218090046197
step: 450, loss: 0.0004957271157763898
step: 460, loss: 0.0001087643249775283
step: 470, loss: 0.0005671812687069178
step: 480, loss: 0.00021496314730029553
step: 490, loss: 0.00024485596804879606
step: 500, loss: 0.0008221847820095718
step: 510, loss: 4.9709204176906496e-05
step: 520, loss: 6.993956776568666e-05
step: 530, loss: 0.0032598599791526794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.


```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9522058823529412, f1=0.9338201734367868,
best_f1=0.9338201734367868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002923830761574209
step: 10, loss: 4.494113818509504e-05
step: 20, loss: 0.00047912204172462225
step: 30, loss: 8.625042391940951e-05
step: 40, loss: 0.00015670373977627605
step: 50, loss: 3.644884782261215e-05
step: 60, loss: 0.00044365739449858665
step: 70, loss: 0.00042751830187626183
step: 80, loss: 5.405253614298999e-05
step: 90, loss: 0.010348938405513763
step: 100, loss: 0.241901695728302
step: 110, loss: 0.0001872439170256257
step: 120, loss: 0.00018530247325543314
step: 130, loss: 0.0006155224400572479
step: 140, loss: 0.0002384747058385983
step: 150, loss: 0.00011408785940147936
step: 160, loss: 0.056557029485702515
step: 170, loss: 0.00019943079678341746
step: 180, loss: 0.00018467854533810169
step: 190, loss: 0.0002072154893539846
step: 200, loss: 0.004901987034827471
step: 210, loss: 0.00021012268553022295
step: 220, loss: 0.00020551065972540528
step: 230, loss: 0.00035584671422839165
step: 240, loss: 0.0011565437307581306
step: 250, loss: 0.0001290489744860679
step: 260, loss: 0.00011652115790639073
step: 270, loss: 0.0018573951674625278
step: 280, loss: 0.00040966051165014505
step: 290, loss: 9.265559492632747e-05
step: 300, loss: 0.001354669569991529
step: 310, loss: 8.915713260648772e-05
step: 320, loss: 0.016642937436699867
step: 330, loss: 0.0018486991757526994
step: 340, loss: 7.095919136190787e-05
step: 350, loss: 0.0002800476213451475
step: 360, loss: 0.0010408245725557208
step: 370, loss: 2.880879765143618e-05

```

```

step: 380, loss: 0.00015964070917107165
step: 390, loss: 0.001539001241326332
step: 400, loss: 0.05587216094136238
step: 410, loss: 0.00010321143781766295
step: 420, loss: 5.738630352425389e-05
step: 430, loss: 3.903930701198988e-05
step: 440, loss: 9.973069245461375e-05
step: 450, loss: 8.751015411689878e-05
step: 460, loss: 4.579768210533075e-05
step: 470, loss: 0.0012372260680422187
step: 480, loss: 8.119390258798376e-05
step: 490, loss: 6.056006168364547e-05
step: 500, loss: 0.00013393483823165298
step: 510, loss: 9.621229401091114e-05
step: 520, loss: 0.00020959821995347738
step: 530, loss: 0.0036693792790174484
/usr/local/lib/python3.7/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=0.951846657316503, f1=0.9402501157943491,
best_f1=0.9338201734367868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.875498704379424e-05
step: 10, loss: 8.232596883317456e-05
step: 20, loss: 0.00044085472472943366
step: 30, loss: 2.9450224246829748e-05
step: 40, loss: 5.02353796036914e-05
step: 50, loss: 0.0005344478995539248
step: 60, loss: 3.729452146217227e-05
step: 70, loss: 6.807444151490927e-05
step: 80, loss: 0.00034736329689621925
step: 90, loss: 5.52220044482965e-05
step: 100, loss: 6.393383227987215e-05
step: 110, loss: 2.1695104805985466e-05
step: 120, loss: 3.13763739541173e-05
step: 130, loss: 6.855648825876415e-05
step: 140, loss: 4.8119210987351835e-05
step: 150, loss: 0.00010356508573750034
step: 160, loss: 0.0007380509050562978
step: 170, loss: 5.196618803893216e-05
step: 180, loss: 3.8401525671361014e-05
step: 190, loss: 0.0004358281148597598
step: 200, loss: 0.09082396328449249
step: 210, loss: 4.245017044013366e-05
step: 220, loss: 2.9551189072662964e-05
step: 230, loss: 0.00018133378762286156

```

```

step: 240, loss: 0.0006861196598038077
step: 250, loss: 8.338066982105374e-05
step: 260, loss: 4.00334938603919e-05
step: 270, loss: 0.000578688457608223
step: 280, loss: 4.836656080442481e-05
step: 290, loss: 3.181404827046208e-05
step: 300, loss: 3.6047465982846916e-05
step: 310, loss: 6.635452882619575e-05
step: 320, loss: 3.5931181628257036e-05
step: 330, loss: 4.7833756980253384e-05
step: 340, loss: 6.0905906138941646e-05
step: 350, loss: 3.140329499728978e-05
step: 360, loss: 0.14318138360977173
step: 370, loss: 4.511345105129294e-05
step: 380, loss: 0.00013572172611020505
step: 390, loss: 0.00016562595556024462
step: 400, loss: 0.0002329127019038424
step: 410, loss: 8.907736628316343e-05
step: 420, loss: 8.653708209749311e-05
step: 430, loss: 0.0001451534335501492
step: 440, loss: 0.00046537129674106836
step: 450, loss: 0.00010194064088864252
step: 460, loss: 0.00029838437330909073
step: 470, loss: 0.0017537453677505255
step: 480, loss: 2.6722387701738626e-05
step: 490, loss: 4.7980462113628164e-05
step: 500, loss: 3.313714842079207e-05
step: 510, loss: 9.563873027218506e-05
step: 520, loss: 4.5469842007150874e-05
step: 530, loss: 5.9904235968133435e-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=0.9503940658321743, f1=0.9348623853211009,
best_f1=0.9338201734367868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013646339066326618
step: 10, loss: 0.004650454968214035
step: 20, loss: 9.433754894416779e-05
step: 30, loss: 6.524235504912212e-05
step: 40, loss: 0.0006133003043942153
step: 50, loss: 9.669832797953859e-05
step: 60, loss: 8.117278048302978e-05
step: 70, loss: 0.00010127606219612062
step: 80, loss: 0.0012954027624800801
step: 90, loss: 4.27869672421366e-05

```

```
step: 100, loss: 9.00767045095563e-05
step: 110, loss: 9.41652397159487e-05
step: 120, loss: 8.949008770287037e-05
step: 130, loss: 0.00013207494339440018
step: 140, loss: 8.936157246353105e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.0007677260437048972
step: 160, loss: 5.5426284234272316e-05
step: 170, loss: 6.527562800329179e-05
step: 180, loss: 5.64095753361471e-05
step: 190, loss: 8.083778811851516e-05
step: 200, loss: 7.730431389063597e-05
step: 210, loss: 0.00036188701051287353
step: 220, loss: 3.085798380197957e-05
step: 230, loss: 5.107977631269023e-05
step: 240, loss: 0.00032304503838531673
step: 250, loss: 0.0002855208585970104
step: 260, loss: 5.863184196641669e-05
step: 270, loss: 0.0001832716225180775
step: 280, loss: 0.00014959636610001326
step: 290, loss: 0.00010624215065035969
step: 300, loss: 0.00015234516467899084
step: 310, loss: 0.11716316640377045
step: 320, loss: 0.00013478480104822665
step: 330, loss: 0.0008169678039848804
step: 340, loss: 0.000839850923512131
step: 350, loss: 8.600085857324302e-05
step: 360, loss: 0.00011243332119192928
step: 370, loss: 0.0006351217161864042
step: 380, loss: 0.007598604075610638
step: 390, loss: 0.00012664451787713915
step: 400, loss: 0.005855197086930275
step: 410, loss: 4.820611502509564e-05
step: 420, loss: 9.628627594793215e-05
step: 430, loss: 0.00010334038233850151
step: 440, loss: 0.00011488874588394538
step: 450, loss: 6.346838199533522e-05
step: 460, loss: 0.00033271132269874215
step: 470, loss: 8.372389856958762e-05
step: 480, loss: 7.0676498580724e-05
step: 490, loss: 0.00010229738109046593
step: 500, loss: 0.0008096711244434118
step: 510, loss: 0.000824293470941484
step: 520, loss: 0.00048016716027632356
step: 530, loss: 0.0027499289717525244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9523368810735771, f1=0.9378166743436204,
best_f1=0.9378166743436204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.987274875631556e-05
step: 10, loss: 0.0011759274639189243
step: 20, loss: 0.0006226691766642034
step: 30, loss: 0.00012735901691485196
step: 40, loss: 6.702105747535825e-05
step: 50, loss: 0.0017370772548019886
step: 60, loss: 9.567289089318365e-05
step: 70, loss: 5.949223123025149e-05
step: 80, loss: 5.0091304728994146e-05
step: 90, loss: 6.294681952567771e-05
step: 100, loss: 3.04470413539093e-05
step: 110, loss: 0.0012027695775032043
step: 120, loss: 4.730835644295439e-05
step: 130, loss: 4.798820737050846e-05
step: 140, loss: 4.3066407670266926e-05
step: 150, loss: 7.971705053932965e-05
step: 160, loss: 7.111068407539278e-05
step: 170, loss: 2.9463617465808056e-05
step: 180, loss: 7.5070362072438e-05
step: 190, loss: 0.0016639773966744542
step: 200, loss: 0.0007522067753598094
step: 210, loss: 2.4771357857389376e-05
step: 220, loss: 5.834191324538551e-05
step: 230, loss: 0.0038220915012061596
step: 240, loss: 5.399970541475341e-05
step: 250, loss: 0.00011421630915720016
step: 260, loss: 3.315576759632677e-05
step: 270, loss: 5.437595973489806e-05
step: 280, loss: 4.734229878522456e-05
step: 290, loss: 5.822563616675325e-05
step: 300, loss: 0.00015303082182072103
step: 310, loss: 0.0001018872790155001
step: 320, loss: 7.650892075616866e-05
step: 330, loss: 6.978552119107917e-05
step: 340, loss: 7.36522997613065e-05
step: 350, loss: 0.0011226864298805594
step: 360, loss: 8.253263513324782e-05
step: 370, loss: 0.00010165785352000967
step: 380, loss: 7.272957736859098e-05

```

```

step: 390, loss: 4.378282756078988e-05
step: 400, loss: 0.0014413567259907722
step: 410, loss: 7.13135814294219e-05
step: 420, loss: 8.99264050531201e-05
step: 430, loss: 2.278585816384293e-05
step: 440, loss: 9.958365262718871e-05
step: 450, loss: 0.0041175405494868755
step: 460, loss: 0.00014048667799215764
step: 470, loss: 5.712547499570064e-05
step: 480, loss: 0.0005371120641939342
step: 490, loss: 5.889406020287424e-05
step: 500, loss: 0.0002714828005991876
step: 510, loss: 6.148598913569003e-05
step: 520, loss: 8.31475990707986e-05
step: 530, loss: 8.44916285132058e-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=0.9522041763341067, f1=0.9372114496768237,
best_f1=0.9378166743436204
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, 160.04it/s]
load_f1 = 0.9503019043195542
real_f1 = 0.9466357308584686
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 135.07it/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: 0.43207401037216187

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=0.2745098039215686,
best_f1=0.2745098039215686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48630404472351074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.29213483146067415, f1=0.23728813559322035,
best_f1=0.23728813559322035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41819339990615845
/usr/local/lib/python3.7/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=0.2916666666666667, f1=0.29508196721311475,
best_f1=0.23728813559322035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29861390590667725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3058823529411765, f1=0.18181818181818182,
best_f1=0.18181818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3425007462501526
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.38095238095238093, f1=0.23076923076923075,
best_f1=0.23076923076923075
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18534091114997864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.43999999999999995, f1=0.6428571428571429,
best_f1=0.6428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3248424530029297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.48, f1=0.5945945945945946, best_f1=0.5945945945945946
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17480316758155823
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 8: dev_f1=0.6470588235294117, f1=0.5517241379310344,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22203126549720764
/usr/local/lib/python3.7/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=0.6451612903225806, f1=0.5925925925925927,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22670526802539825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7857142857142857, f1=0.6153846153846153,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1504192054271698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8125000000000001, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10314523428678513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.689655172413793,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.034026551991701126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8571428571428571, f1=0.689655172413793,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04310120269656181
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.689655172413793,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03754141554236412
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.689655172413793,
best_f1=0.689655172413793
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, 119874.89it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.631578947368421
real_f1 = 0.6190476190476191
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 135.47it/s]

```

4.1.2 DBLP-ACM - Running the matcher

```

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

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

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

```

```

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 398kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 745kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 504kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 68.2MB/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: 0.5694915652275085
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.4613169729709625
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: 0.507539689540863
step: 30, loss: 0.21069364249706268
step: 40, loss: 0.10192585736513138
step: 50, loss: 0.12339179217815399
step: 60, loss: 0.1507759690284729
step: 70, loss: 0.03984959051012993

```

```

step: 80, loss: 0.009184087626636028
step: 90, loss: 0.1408015489578247
step: 100, loss: 0.13787153363227844
step: 110, loss: 0.03728071227669716
step: 120, loss: 0.04943274334073067
step: 130, loss: 0.0033714494202286005
step: 140, loss: 0.11942315846681595
step: 150, loss: 0.049290772527456284
step: 160, loss: 0.0023574314545840025
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.15831008553504944
step: 180, loss: 0.011704863049089909
step: 190, loss: 0.17506876587867737
step: 200, loss: 0.013803992420434952
step: 210, loss: 0.15868806838989258
step: 220, loss: 0.057718031108379364
step: 230, loss: 0.0022640316747128963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9743589743589743, f1=0.979591836734694,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022502713836729527
step: 10, loss: 0.15093767642974854
step: 20, loss: 0.005951005499809998
step: 30, loss: 0.07108789682388306
step: 40, loss: 0.05230899900197983
step: 50, loss: 0.009586113505065441
step: 60, loss: 0.01238249707967043
step: 70, loss: 0.010657447390258312
step: 80, loss: 0.002654158743098378
step: 90, loss: 0.01016421988606453
step: 100, loss: 0.002609098795801401
step: 110, loss: 0.002627011388540268
step: 120, loss: 0.07292088121175766
step: 130, loss: 0.006501822266727686
step: 140, loss: 0.0020083419512957335
step: 150, loss: 0.054591692984104156
step: 160, loss: 0.0019229196477681398
step: 170, loss: 0.0007623211131431162
step: 180, loss: 0.009562785737216473

```

```

step: 190, loss: 0.0017743132775649428
step: 200, loss: 0.007869131863117218
step: 210, loss: 0.02951963059604168
step: 220, loss: 0.000314743141643703
step: 230, loss: 0.0015910138608887792
/usr/local/lib/python3.7/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=0.9672316384180792, f1=0.9705215419501134,
best_f1=0.979591836734694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017738018650561571
step: 10, loss: 0.1445198655128479
step: 20, loss: 0.02079848013818264
step: 30, loss: 0.00152308598626405
step: 40, loss: 0.013963916338980198
step: 50, loss: 0.0908055305480957
step: 60, loss: 0.0030108068604022264
step: 70, loss: 0.0017682993784546852
step: 80, loss: 0.0455356128513813
step: 90, loss: 0.0007067496189847589
step: 100, loss: 0.0013472149148583412
step: 110, loss: 0.0012834100052714348
step: 120, loss: 0.003845105180516839
step: 130, loss: 0.00270848092623055
step: 140, loss: 0.002570709679275751
step: 150, loss: 0.009603319689631462
step: 160, loss: 0.0016790678491815925
step: 170, loss: 0.0010168240405619144
step: 180, loss: 0.00691266031935811
step: 190, loss: 0.01516785193234682
step: 200, loss: 0.041997842490673065
step: 210, loss: 0.0018605379154905677
step: 220, loss: 0.00804523378610611
step: 230, loss: 0.01937943696975708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.983050847457627, f1=0.9782857142857143,
best_f1=0.9782857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006432270165532827
step: 10, loss: 0.014370436780154705
step: 20, loss: 0.03250017762184143
step: 30, loss: 0.0005452783079817891
step: 40, loss: 0.0006902352324686944
step: 50, loss: 0.03703226149082184
step: 60, loss: 0.002944356994703412
step: 70, loss: 0.001936351298354566
step: 80, loss: 0.0002704510116018355
step: 90, loss: 0.005270557478070259
step: 100, loss: 0.002450656145811081
step: 110, loss: 0.0010409998940303922
step: 120, loss: 0.0007167558069340885
step: 130, loss: 0.004830727819353342
step: 140, loss: 0.003644038923084736
step: 150, loss: 0.024808522313833237
step: 160, loss: 0.0028994339518249035
step: 170, loss: 0.011192851699888706
step: 180, loss: 0.19180767238140106
step: 190, loss: 0.016708463430404663
step: 200, loss: 0.01250758022069931
step: 210, loss: 0.0009586401283740997
step: 220, loss: 0.00721599068492651
step: 230, loss: 0.006919607520103455
/usr/local/lib/python3.7/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=0.9828962371721778, f1=0.9759450171821306,
best_f1=0.9782857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004664764739573002
step: 10, loss: 0.0026731400284916162
step: 20, loss: 0.0016041317721828818
step: 30, loss: 0.0004350129747763276
step: 40, loss: 0.0008097626268863678
step: 50, loss: 0.0008459163946099579
step: 60, loss: 0.007708555553108454
step: 70, loss: 0.0015971404500305653
step: 80, loss: 0.06798835843801498
step: 90, loss: 0.0678967610001564
step: 100, loss: 0.0004653066862374544
step: 110, loss: 0.001554856076836586
step: 120, loss: 0.00032402490614913404
step: 130, loss: 0.01663735881447792
step: 140, loss: 0.00039615685818716884

```

```

step: 150, loss: 0.0313178151845932
step: 160, loss: 0.0011719001922756433
step: 170, loss: 0.0007589748711325228
step: 180, loss: 0.0005692386766895652
step: 190, loss: 0.0005768210394307971
step: 200, loss: 0.04995523393154144
step: 210, loss: 0.051226209849119186
step: 220, loss: 0.0006408441695384681
step: 230, loss: 0.004609519615769386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853438556933484, f1=0.9852440408626559,
best_f1=0.9852440408626559
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003009194042533636
step: 10, loss: 0.0011953309876844287
step: 20, loss: 0.0014507743762806058
step: 30, loss: 0.0007055288879200816
step: 40, loss: 0.000321993138641119
step: 50, loss: 0.0029765008948743343
step: 60, loss: 0.0075348783284425735
step: 70, loss: 0.0032271852251142263
step: 80, loss: 0.0009091655374504626
step: 90, loss: 0.00588660454377532
step: 100, loss: 0.0032878268975764513
step: 110, loss: 0.0138554647564888
step: 120, loss: 0.0003563401405699551
step: 130, loss: 0.003484228625893593
step: 140, loss: 0.0006283211405389011
step: 150, loss: 0.0005776354810222983
step: 160, loss: 0.0287152249366045
step: 170, loss: 0.0027523718308657408
step: 180, loss: 9.067869541468099e-05
step: 190, loss: 0.00026656800764612854
step: 200, loss: 0.020388148725032806
step: 210, loss: 0.02328602410852909
step: 220, loss: 0.03387392312288284
step: 230, loss: 0.0034003991167992353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853768278965129, f1=0.9841628959276018,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016626982251182199
step: 10, loss: 0.0004479703784454614
step: 20, loss: 0.0008143963059410453
step: 30, loss: 0.001710855751298368
step: 40, loss: 0.020338337868452072
step: 50, loss: 0.004768391605466604
step: 60, loss: 0.0012602602364495397
step: 70, loss: 0.0019362817984074354
step: 80, loss: 0.0011020818492397666
step: 90, loss: 0.020069967955350876
step: 100, loss: 0.00023465807316824794
step: 110, loss: 0.0010817135917022824
step: 120, loss: 0.028427261859178543
step: 130, loss: 0.0046137250028550625
step: 140, loss: 0.0002380896476097405
step: 150, loss: 0.04187675192952156
step: 160, loss: 0.0001810515095712617
step: 170, loss: 0.028932882472872734
step: 180, loss: 0.0011870766757056117
step: 190, loss: 0.007633597124367952
step: 200, loss: 0.009779025800526142
step: 210, loss: 0.10584932565689087
step: 220, loss: 0.00023663636238779873
step: 230, loss: 0.0005730170523747802
/usr/local/lib/python3.7/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=0.984090909090909, f1=0.9807474518686297,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005363897071219981
step: 10, loss: 0.01955900900065899
step: 20, loss: 0.0004288141499273479
step: 30, loss: 0.00035757457953877747
step: 40, loss: 0.0028571265283972025
step: 50, loss: 0.001139925909228623
step: 60, loss: 0.00034536735620349646

```

```

step: 70, loss: 0.00016576026973780245
step: 80, loss: 0.020123867318034172
step: 90, loss: 0.001693632686510682
step: 100, loss: 0.00034452060936018825
step: 110, loss: 0.004175763111561537
step: 120, loss: 0.0006478888099081814
step: 130, loss: 0.0017193349776789546
step: 140, loss: 0.0005979957059025764
step: 150, loss: 0.08098382502794266
step: 160, loss: 0.007891025394201279
step: 170, loss: 0.011208636686205864
step: 180, loss: 0.00029945725691504776
step: 190, loss: 0.0009459995198994875
step: 200, loss: 0.0024857677053660154
step: 210, loss: 0.005779838189482689
step: 220, loss: 0.00022472592536360025
step: 230, loss: 8.504246216034517e-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=0.9842696629213483, f1=0.9807909604519773,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001998286898015067
step: 10, loss: 0.0012532470282167196
step: 20, loss: 0.0001818884047679603
step: 30, loss: 0.0012933035613968968
step: 40, loss: 0.0025668616872280836
step: 50, loss: 0.011644613929092884
step: 60, loss: 0.0014631320955231786
step: 70, loss: 0.035567834973335266
step: 80, loss: 0.00869214441627264
step: 90, loss: 0.03323604539036751
step: 100, loss: 0.00018975819693878293
step: 110, loss: 0.00015430210623890162
step: 120, loss: 0.03501712903380394
step: 130, loss: 0.0002469326718710363
step: 140, loss: 0.00492213387042284
step: 150, loss: 0.00021205018856562674
step: 160, loss: 0.00043325175647623837
step: 170, loss: 0.00017511684563942254
step: 180, loss: 0.0001175965735455975
step: 190, loss: 0.0004888399271294475
step: 200, loss: 0.000900097016710788
step: 210, loss: 0.0049918945878744125
step: 220, loss: 0.0002360124053666368

```

```

step: 230, loss: 0.008025788702070713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9875424688561721, f1=0.9783845278725825,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003638937370851636
step: 10, loss: 0.0008248501690104604
step: 20, loss: 0.00029974390054121614
step: 30, loss: 9.485911868978292e-05
step: 40, loss: 0.049444183707237244
step: 50, loss: 0.011743253096938133
step: 60, loss: 8.110709313768893e-05
step: 70, loss: 0.0003236409684177488
step: 80, loss: 0.00010121728701051325
step: 90, loss: 0.000272795936325565
step: 100, loss: 0.00014726108929608017
step: 110, loss: 0.0037405395414680243
step: 120, loss: 7.096554327290505e-05
step: 130, loss: 0.00011270320828771219
step: 140, loss: 0.0001072974264388904
step: 150, loss: 0.0005142163718119264
step: 160, loss: 0.0013226292794570327
step: 170, loss: 0.00013878862955607474
step: 180, loss: 0.0010031687561422586
step: 190, loss: 0.002458944683894515
step: 200, loss: 0.0004864686634391546
step: 210, loss: 0.0020726146176457405
step: 220, loss: 0.0033794178161770105
step: 230, loss: 0.00017139124975074083
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9853438556933484,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011860465019708499
step: 10, loss: 0.00012914507533423603
step: 20, loss: 0.00019757109112106264

```

```

step: 30, loss: 0.0006777283269912004
step: 40, loss: 0.0008123057777993381
step: 50, loss: 0.0004929104470647871
step: 60, loss: 0.006315091159194708
step: 70, loss: 0.15796814858913422
step: 80, loss: 0.019466308876872063
step: 90, loss: 0.08412351459264755
step: 100, loss: 0.005442132707685232
step: 110, loss: 0.013899503275752068
step: 120, loss: 0.0004095017211511731
step: 130, loss: 0.00024817293160595
step: 140, loss: 0.000636428885627538
step: 150, loss: 0.0011013188632205129
step: 160, loss: 0.042892176657915115
step: 170, loss: 0.0002780552313197404
step: 180, loss: 0.00014605297474190593
step: 190, loss: 0.0002954913070425391
step: 200, loss: 0.004977263044565916
step: 210, loss: 0.0002083014405798167
step: 220, loss: 0.0011995803797617555
step: 230, loss: 0.00015631182759534568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9865771812080537, f1=0.9798206278026906,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001172863645479083
step: 10, loss: 0.00011832373274955899
step: 20, loss: 0.0068227252922952175
step: 30, loss: 0.11596833914518356
step: 40, loss: 0.00012345437426120043
step: 50, loss: 0.0008707887609489262
step: 60, loss: 0.00020220484293531626
step: 70, loss: 0.00034803844755515456
step: 80, loss: 0.0001341018360108137
step: 90, loss: 0.0018143391935154796
step: 100, loss: 8.913539932109416e-05
step: 110, loss: 0.00016934028826653957
step: 120, loss: 0.00015117487055249512
step: 130, loss: 0.00010886645031860098
step: 140, loss: 8.404238178627566e-05
step: 150, loss: 0.0007400392787531018
step: 160, loss: 0.0002505868033040315
step: 170, loss: 0.0005075481021776795
step: 180, loss: 7.555662887170911e-05

```

```

step: 190, loss: 0.00019890637486241758
step: 200, loss: 0.00025124027160927653
step: 210, loss: 0.0002735171001404524
step: 220, loss: 0.0006083884509280324
step: 230, loss: 0.00020209577633067966
/usr/local/lib/python3.7/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=0.9865771812080537, f1=0.9776785714285714,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012827612226828933
step: 10, loss: 0.00010379629384260625
step: 20, loss: 0.00013892391871195287
step: 30, loss: 7.61329720262438e-05
step: 40, loss: 0.0005943099386058748
step: 50, loss: 0.0002892198972404003
step: 60, loss: 9.108435187954456e-05
step: 70, loss: 0.0001320799347013235
step: 80, loss: 8.414711192017421e-05
step: 90, loss: 0.00026652473025023937
step: 100, loss: 0.0002559784916229546
step: 110, loss: 0.00011648366489680484
step: 120, loss: 0.001258753938600421
step: 130, loss: 7.248381734825671e-05
step: 140, loss: 0.00038712978130206466
step: 150, loss: 0.0232806745916605
step: 160, loss: 0.00025608859141357243
step: 170, loss: 9.503385808784515e-05
step: 180, loss: 0.023506058380007744
step: 190, loss: 0.0001828851964091882
step: 200, loss: 3.9914153603604063e-05
step: 210, loss: 0.0018980479799211025
step: 220, loss: 0.000814126746263355
step: 230, loss: 0.00031494026188738644
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9820627802690582,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.984079714631662e-05
step: 10, loss: 7.370977255050093e-05
step: 20, loss: 0.00026892084861174226

```

```

step: 30, loss: 0.0001179735190817155
step: 40, loss: 0.00013109839346725494
step: 50, loss: 4.7017030738061294e-05
step: 60, loss: 9.863166633294895e-05
step: 70, loss: 7.066014950396493e-05
step: 80, loss: 5.085350858280435e-05
step: 90, loss: 8.228201477322727e-05
step: 100, loss: 8.351852738996968e-05
step: 110, loss: 0.00016536563634872437
step: 120, loss: 0.0002387937856838107
step: 130, loss: 9.081680764211342e-05
step: 140, loss: 0.00011093800276285037
step: 150, loss: 6.789807957829908e-05
step: 160, loss: 7.287931657629088e-05
step: 170, loss: 0.00016110904107335955
step: 180, loss: 8.811177394818515e-05
step: 190, loss: 0.0005318194744177163
step: 200, loss: 0.00040860011358745396
step: 210, loss: 0.0001469938870286569
step: 220, loss: 0.0001719787687761709
step: 230, loss: 7.157767686294392e-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=0.9865771812080537, f1=0.977728285077951,
best_f1=0.9783845278725825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01566254533827305
step: 10, loss: 6.078672595322132e-05
step: 20, loss: 0.00040293586789630353
step: 30, loss: 0.0005762849468737841
step: 40, loss: 0.00021294064936228096
step: 50, loss: 5.149361822986975e-05
step: 60, loss: 0.022717667743563652
step: 70, loss: 0.00017823050438892096
step: 80, loss: 0.0003068339719902724
step: 90, loss: 5.820422666147351e-05
step: 100, loss: 3.900482988683507e-05
step: 110, loss: 8.452699694316834e-05
step: 120, loss: 0.025097601115703583
step: 130, loss: 8.256967703346163e-05
step: 140, loss: 0.022570330649614334
step: 150, loss: 0.00017993392248172313
step: 160, loss: 0.003323034616187215
step: 170, loss: 7.802828622516245e-05
step: 180, loss: 0.016489550471305847

```



```

step: 190, loss: 0.010693062096834183
step: 200, loss: 0.00019365563639439642
step: 210, loss: 0.11426997184753418
step: 220, loss: 8.275217987829819e-05
step: 230, loss: 0.00013292775838635862
/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9788182831661093,
best_f1=0.9783845278725825
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:16, 150.41it/s]
load_f1 = 0.9853438556933484
real_f1 = 0.9853438556933484
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, 138.75it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale              : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.606137216091156
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: 0.3813115358352661
step: 20, loss: 0.3339666426181793
step: 30, loss: 0.34618377685546875
step: 40, loss: 0.29740557074546814
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.26631200313568115
step: 60, loss: 0.2901190221309662
step: 70, loss: 0.21533145010471344
step: 80, loss: 0.21680372953414917
step: 90, loss: 0.2564099133014679
step: 100, loss: 0.1649821698665619
step: 110, loss: 0.1858101487159729
step: 120, loss: 0.41602399945259094
step: 130, loss: 0.15681520104408264
step: 140, loss: 0.20255662500858307
step: 150, loss: 0.1738733947277069
step: 160, loss: 0.17416957020759583
step: 170, loss: 0.12701022624969482
step: 180, loss: 0.10339148342609406
step: 190, loss: 0.21306608617305756
step: 200, loss: 0.06310432404279709
step: 210, loss: 0.03468667343258858
step: 220, loss: 0.1239701583981514
step: 230, loss: 0.30094337463378906
step: 240, loss: 0.08727677166461945
step: 250, loss: 0.05579999089241028
step: 260, loss: 0.1561049073934555
step: 270, loss: 0.3873644769191742
step: 280, loss: 0.04489349201321602
step: 290, loss: 0.07782834023237228

```

```

step: 300, loss: 0.07372915744781494
step: 310, loss: 0.4715902805328369
step: 320, loss: 0.07925978302955627
step: 330, loss: 0.06272384524345398
step: 340, loss: 0.30843979120254517
step: 350, loss: 0.13164041936397552
step: 360, loss: 0.00876579713076353
step: 370, loss: 0.005978443659842014
step: 380, loss: 0.18496258556842804
step: 390, loss: 0.02200998179614544
step: 400, loss: 0.08904711157083511
step: 410, loss: 0.2905789315700531
step: 420, loss: 0.028638944029808044
step: 430, loss: 0.05943340063095093
step: 440, loss: 0.01543080247938633
step: 450, loss: 0.13209523260593414
step: 460, loss: 0.07415017485618591
step: 470, loss: 0.062124721705913544
step: 480, loss: 0.1052551120519638
step: 490, loss: 0.20712143182754517
step: 500, loss: 0.22502431273460388
step: 510, loss: 0.15266625583171844
step: 520, loss: 0.16174015402793884
step: 530, loss: 0.10321137309074402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9208962048468221, f1=0.9241942805265547,
best_f1=0.9241942805265547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13317248225212097
step: 10, loss: 0.05261015519499779
step: 20, loss: 0.026897408068180084
step: 30, loss: 0.04680647328495979
step: 40, loss: 0.03810486942529678
step: 50, loss: 0.05564562976360321
step: 60, loss: 0.013887295499444008
step: 70, loss: 0.016162976622581482
step: 80, loss: 0.029929231852293015
step: 90, loss: 0.01784241385757923
step: 100, loss: 0.2047959417104721
step: 110, loss: 0.019473344087600708

```

```

step: 120, loss: 0.17577141523361206
step: 130, loss: 0.003411678597331047
step: 140, loss: 0.2796936333179474
step: 150, loss: 0.019736764952540398
step: 160, loss: 0.07065130770206451
step: 170, loss: 0.07055088132619858
step: 180, loss: 0.023690907284617424
step: 190, loss: 0.022595256567001343
step: 200, loss: 0.2100331038236618
step: 210, loss: 0.03379223868250847
step: 220, loss: 0.0025464606005698442
step: 230, loss: 0.039422657340765
step: 240, loss: 0.11287260800600052
step: 250, loss: 0.019246544688940048
step: 260, loss: 0.12159252166748047
step: 270, loss: 0.13476449251174927
step: 280, loss: 0.07590527087450027
step: 290, loss: 0.05892631411552429
step: 300, loss: 0.026032704859972
step: 310, loss: 0.05722814053297043
step: 320, loss: 0.07739106565713882
step: 330, loss: 0.052425138652324677
step: 340, loss: 0.12815186381340027
step: 350, loss: 0.0034228800795972347
step: 360, loss: 0.11864417046308517
step: 370, loss: 0.022051004692912102
step: 380, loss: 0.10011057555675507
step: 390, loss: 0.006891404744237661
step: 400, loss: 0.04820229858160019
step: 410, loss: 0.06843344867229462
step: 420, loss: 0.11726196855306625
step: 430, loss: 0.10012684762477875
step: 440, loss: 0.008054491132497787
step: 450, loss: 0.0628364160656929
step: 460, loss: 0.03143753856420517
step: 470, loss: 0.03738342598080635
step: 480, loss: 0.10280831903219223
step: 490, loss: 0.06577471643686295
step: 500, loss: 0.026015792042016983
step: 510, loss: 0.03544866666197777
step: 520, loss: 0.37374669313430786
step: 530, loss: 0.03510481119155884
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9396035039188566, f1=0.9382488479262673,
best_f1=0.9382488479262673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07316377013921738
step: 10, loss: 0.04281625896692276
step: 20, loss: 0.01631820946931839
step: 30, loss: 0.08127475529909134
step: 40, loss: 0.04697435349225998
step: 50, loss: 0.010258747264742851
step: 60, loss: 0.04668303206562996
step: 70, loss: 0.005345093552023172
step: 80, loss: 0.10208022594451904
step: 90, loss: 0.053034424781799316
step: 100, loss: 0.046690817922353745
step: 110, loss: 0.03751806914806366
step: 120, loss: 0.05420231446623802
step: 130, loss: 0.06194634735584259
step: 140, loss: 0.020359523594379425
step: 150, loss: 0.04749861732125282
step: 160, loss: 0.04622239992022514
step: 170, loss: 0.01747662015259266
step: 180, loss: 0.012719821184873581
step: 190, loss: 0.007900817319750786
step: 200, loss: 0.02136331982910633
step: 210, loss: 0.08450683206319809
step: 220, loss: 0.1682722568511963
step: 230, loss: 0.08784327656030655
step: 240, loss: 0.06080523505806923
step: 250, loss: 0.1537419855594635
step: 260, loss: 0.18353396654129028
step: 270, loss: 0.012953092344105244
step: 280, loss: 0.03796635568141937
step: 290, loss: 0.06130484491586685
step: 300, loss: 0.14984072744846344
step: 310, loss: 0.0789572224020958
step: 320, loss: 0.07684384286403656
step: 330, loss: 0.022490402683615685
step: 340, loss: 0.07074639201164246
step: 350, loss: 0.2649002969264984
step: 360, loss: 0.04028891399502754
step: 370, loss: 0.10721617937088013
step: 380, loss: 0.011283115483820438
step: 390, loss: 0.007688898127526045
step: 400, loss: 0.1757425218820572
step: 410, loss: 0.022735584527254105
```

```

step: 420, loss: 0.010358680970966816
step: 430, loss: 0.04032671079039574
step: 440, loss: 0.26247814297676086
step: 450, loss: 0.03932075947523117
step: 460, loss: 0.11997988820075989
step: 470, loss: 0.06252006441354752
step: 480, loss: 0.08338947594165802
step: 490, loss: 0.0070066386833786964
step: 500, loss: 0.018661944195628166
step: 510, loss: 0.14007945358753204
step: 520, loss: 0.04873284697532654
step: 530, loss: 0.010572012513875961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9416745061147694, f1=0.9317004239284032,
best_f1=0.9317004239284032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022154806181788445
step: 10, loss: 0.01835184544324875
step: 20, loss: 0.17791423201560974
step: 30, loss: 0.12820959091186523
step: 40, loss: 0.04535778984427452
step: 50, loss: 0.06600817292928696
step: 60, loss: 0.00863068550825119
step: 70, loss: 0.024742990732192993
step: 80, loss: 0.14743158221244812
step: 90, loss: 0.029787929728627205
step: 100, loss: 0.04672364890575409
step: 110, loss: 0.08934522420167923
step: 120, loss: 0.006670564413070679
step: 130, loss: 0.021267782896757126
step: 140, loss: 0.16181395947933197
step: 150, loss: 0.013502643443644047
step: 160, loss: 0.006910014897584915
step: 170, loss: 0.048135433346033096
step: 180, loss: 0.07782328873872757
step: 190, loss: 0.0760607123374939
step: 200, loss: 0.1124832034111023
step: 210, loss: 0.0065460121259093285
step: 220, loss: 0.03736262023448944
step: 230, loss: 0.006350819021463394

```

```

step: 240, loss: 0.008487462997436523
step: 250, loss: 0.19775064289569855
step: 260, loss: 0.0013966687256470323
step: 270, loss: 0.06365086138248444
step: 280, loss: 0.0011491129407659173
step: 290, loss: 0.007872110232710838
step: 300, loss: 0.007234976626932621
step: 310, loss: 0.0029694801196455956
step: 320, loss: 0.07045415043830872
step: 330, loss: 0.008470803499221802
step: 340, loss: 0.019922126084566116
step: 350, loss: 0.16380469501018524
step: 360, loss: 0.04913759604096413
step: 370, loss: 0.003966163843870163
step: 380, loss: 0.003832476679235697
step: 390, loss: 0.003697273787111044
step: 400, loss: 0.04776689037680626
step: 410, loss: 0.0035624064039438963
step: 420, loss: 0.007249858230352402
step: 430, loss: 0.011952937580645084
step: 440, loss: 0.025097176432609558
step: 450, loss: 0.020912615582346916
step: 460, loss: 0.023274948820471764
step: 470, loss: 0.007091139443218708
step: 480, loss: 0.009568140842020512
step: 490, loss: 0.023100553080439568
step: 500, loss: 0.03412175178527832
step: 510, loss: 0.141281396150589
step: 520, loss: 0.03086145967245102
step: 530, loss: 0.06278584152460098
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9466357308584686, f1=0.9471733086190918,
best_f1=0.9471733086190918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014932365156710148
step: 10, loss: 0.005031618755310774
step: 20, loss: 0.0054555367678403854
step: 30, loss: 0.021153060719370842
step: 40, loss: 0.012434188276529312
step: 50, loss: 0.005781815852969885

```


step: 60, loss: 0.12940531969070435
step: 70, loss: 0.0032514366321265697
step: 80, loss: 0.0026372119318693876
step: 90, loss: 0.01799567975103855
step: 100, loss: 0.01084202155470848
step: 110, loss: 0.011569601483643055
step: 120, loss: 0.0998581275343895
step: 130, loss: 0.004069410730153322
step: 140, loss: 0.06353927403688431
step: 150, loss: 0.014125900343060493
step: 160, loss: 0.03478146344423294
step: 170, loss: 0.07610362768173218
step: 180, loss: 0.003680173773318529
step: 190, loss: 0.006741248071193695
step: 200, loss: 0.020960986614227295
step: 210, loss: 0.014063806273043156
step: 220, loss: 0.02724333666265011
step: 230, loss: 0.0025621610693633556
step: 240, loss: 0.002139428863301873
step: 250, loss: 0.13625997304916382
step: 260, loss: 0.0034429491497576237
step: 270, loss: 0.0021824820432811975
step: 280, loss: 0.00470541650429368
step: 290, loss: 0.0008850482990965247
step: 300, loss: 0.07385066896677017
step: 310, loss: 0.0814322680234909
step: 320, loss: 0.0834287479519844
step: 330, loss: 0.0072808717377483845
step: 340, loss: 0.005370112136006355
step: 350, loss: 0.0020600873976945877
step: 360, loss: 0.0036305496469140053
step: 370, loss: 0.001188992871902883
step: 380, loss: 0.00078268360812217
step: 390, loss: 0.0028429620433598757
step: 400, loss: 0.014043509028851986
step: 410, loss: 0.10081568360328674
step: 420, loss: 0.1764100044965744
step: 430, loss: 0.15545393526554108
step: 440, loss: 0.0028862745966762304
step: 450, loss: 0.0744299665093422
step: 460, loss: 0.06024523451924324
step: 470, loss: 0.08876417577266693
step: 480, loss: 0.05068576708436012
step: 490, loss: 0.043092142790555954
step: 500, loss: 0.05440157279372215
step: 510, loss: 0.005979736801236868
step: 520, loss: 0.031096097081899643
step: 530, loss: 0.00821936409920454

```
/usr/local/lib/python3.7/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=0.9443929564411491, f1=0.9377593360995851,
best_f1=0.9471733086190918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0058310073800385
step: 10, loss: 0.016502194106578827
step: 20, loss: 0.013568859547376633
step: 30, loss: 0.0005879816017113626
step: 40, loss: 0.0009337028604932129
step: 50, loss: 0.0025594234466552734
step: 60, loss: 0.006940608378499746
step: 70, loss: 0.023079222068190575
step: 80, loss: 0.00030531344236806035
step: 90, loss: 0.002775172470137477
step: 100, loss: 0.07166601717472076
step: 110, loss: 0.0151442252099514
step: 120, loss: 0.027859238907694817
step: 130, loss: 0.0024020280689001083
step: 140, loss: 0.0009936299175024033
step: 150, loss: 0.0008103420841507614
step: 160, loss: 0.08162069320678711
step: 170, loss: 0.008264065720140934
step: 180, loss: 0.04255475476384163
step: 190, loss: 0.19601735472679138
step: 200, loss: 0.012622278183698654
step: 210, loss: 0.004113473929464817
step: 220, loss: 0.019340183585882187
step: 230, loss: 0.006378346122801304
step: 240, loss: 0.02351589873433113
step: 250, loss: 0.09315440058708191
step: 260, loss: 0.0008353099110536277
step: 270, loss: 0.005912075750529766
step: 280, loss: 0.08176819980144501
step: 290, loss: 0.0037846534978598356
step: 300, loss: 0.018126370385289192
step: 310, loss: 0.039399489760398865
step: 320, loss: 0.0011082241544499993
step: 330, loss: 0.005292197223752737
step: 340, loss: 0.00034680782118812203
step: 350, loss: 0.005121369380503893
step: 360, loss: 0.03772885724902153
step: 370, loss: 0.01914963871240616
step: 380, loss: 0.002284039743244648
step: 390, loss: 0.00233170366846025
```

```

step: 400, loss: 0.005287799518555403
step: 410, loss: 0.00023576797684654593
step: 420, loss: 0.007376644294708967
step: 430, loss: 0.010490200482308865
step: 440, loss: 0.054601363837718964
step: 450, loss: 0.25887465476989746
step: 460, loss: 0.0028668621089309454
step: 470, loss: 0.008637366816401482
step: 480, loss: 0.006758435629308224
step: 490, loss: 0.00469678221270442
step: 500, loss: 0.03657536953687668
step: 510, loss: 0.0055360132828354836
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 520, loss: 0.00043652861495502293
step: 530, loss: 0.011150254867970943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9487418452935694, f1=0.9399720800372265,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014206666965037584
step: 10, loss: 0.01293089147657156
step: 20, loss: 0.0026962950360029936
step: 30, loss: 0.021569784730672836
step: 40, loss: 0.07341066747903824
step: 50, loss: 0.010645671747624874
step: 60, loss: 0.01602829061448574
step: 70, loss: 0.022798730060458183
step: 80, loss: 0.0011981852585449815
step: 90, loss: 0.0014885213458910584
step: 100, loss: 0.00836638268083334
step: 110, loss: 0.001604898483492434
step: 120, loss: 0.0012592682614922523
step: 130, loss: 0.0027831955812871456
step: 140, loss: 0.004111242946237326
step: 150, loss: 0.004650472197681665
step: 160, loss: 0.0004686176835093647
step: 170, loss: 0.0041988203302025795
step: 180, loss: 0.17467302083969116
step: 190, loss: 0.11747804284095764
step: 200, loss: 0.0003316267975606024

```

```

step: 210, loss: 0.0011544371955096722
step: 220, loss: 0.005992097780108452
step: 230, loss: 0.002538035623729229
step: 240, loss: 0.04585814103484154
step: 250, loss: 0.0055759078823029995
step: 260, loss: 0.0008416881901212037
step: 270, loss: 0.02098550647497177
step: 280, loss: 0.0062393732368946075
step: 290, loss: 0.012894675135612488
step: 300, loss: 0.0004961987142451108
step: 310, loss: 0.002068087924271822
step: 320, loss: 0.005859371740370989
step: 330, loss: 0.020264193415641785
step: 340, loss: 0.007309906184673309
step: 350, loss: 0.005239480175077915
step: 360, loss: 0.007459644693881273
step: 370, loss: 0.031967610120773315
step: 380, loss: 0.022208090871572495
step: 390, loss: 0.028261195868253708
step: 400, loss: 0.08877413719892502
step: 410, loss: 0.03677643463015556
step: 420, loss: 0.012775941751897335
step: 430, loss: 0.0007754291291348636
step: 440, loss: 0.004089847672730684
step: 450, loss: 0.00850016251206398
step: 460, loss: 0.02185203693807125
step: 470, loss: 0.13748152554035187
step: 480, loss: 0.005674354266375303
step: 490, loss: 0.010912875644862652
step: 500, loss: 0.013223385438323021
step: 510, loss: 0.0021013759542256594
step: 520, loss: 0.0010090019786730409
step: 530, loss: 0.0019997877534478903
/usr/local/lib/python3.7/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=0.9421028253821214, f1=0.9377901578458682,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002069563139230013
step: 10, loss: 0.0019284005975350738
step: 20, loss: 0.0011496013030409813
step: 30, loss: 0.0006752717890776694
step: 40, loss: 0.00022162377717904747
step: 50, loss: 0.00044391624396666884
step: 60, loss: 0.003026674734428525

```

step: 70, loss: 0.0024018683470785618
step: 80, loss: 0.043764613568782806
step: 90, loss: 0.0009156158776022494
step: 100, loss: 0.00162324879784137
step: 110, loss: 0.0044313655234873295
step: 120, loss: 0.0004778243019245565
step: 130, loss: 0.0009122313931584358
step: 140, loss: 0.03972174599766731
step: 150, loss: 0.0009108323720283806
step: 160, loss: 0.00010731707880040631
step: 170, loss: 0.15172021090984344
step: 180, loss: 0.002429247833788395
step: 190, loss: 0.04995483160018921
step: 200, loss: 0.009895405732095242
step: 210, loss: 0.18221309781074524
step: 220, loss: 0.0015077907592058182
step: 230, loss: 0.008135531097650528
step: 240, loss: 0.03179322928190231
step: 250, loss: 0.00561326602473855
step: 260, loss: 0.0004110360750928521
step: 270, loss: 0.020417237654328346
step: 280, loss: 0.00024149577075149864
step: 290, loss: 0.03798707202076912
step: 300, loss: 2.8735372325172648e-05
step: 310, loss: 0.0018658485496416688
step: 320, loss: 0.00015113467816263437
step: 330, loss: 0.0001241179706994444
step: 340, loss: 0.005048224236816168
step: 350, loss: 6.645944085903466e-05
step: 360, loss: 0.20523115992546082
step: 370, loss: 0.1413818597793579
step: 380, loss: 0.001422283472493291
step: 390, loss: 0.029978666454553604
step: 400, loss: 0.007643405348062515
step: 410, loss: 0.0011241957545280457
step: 420, loss: 0.0016274229856207967
step: 430, loss: 0.001807165564969182
step: 440, loss: 0.14102846384048462
step: 450, loss: 0.0011914579663425684
step: 460, loss: 0.005523762200027704
step: 470, loss: 0.12594151496887207
step: 480, loss: 0.010330304503440857
step: 490, loss: 0.002976295305415988
step: 500, loss: 0.0029989504255354404
step: 510, loss: 0.007789120078086853
step: 520, loss: 0.0028056404553353786
step: 530, loss: 0.004769053775817156
/usr/local/lib/python3.7/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=0.943758573388203, f1=0.9419237749546279,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006963172345422208
step: 10, loss: 0.002032934920862317
step: 20, loss: 0.008060493506491184
step: 30, loss: 0.06687192618846893
step: 40, loss: 0.0005011147004552186
step: 50, loss: 0.0016683556605130434
step: 60, loss: 0.00035380979534238577
step: 70, loss: 0.015431622974574566
step: 80, loss: 0.01239364966750145
step: 90, loss: 0.10816390812397003
step: 100, loss: 0.0008125798194669187
step: 110, loss: 0.029460521414875984
step: 120, loss: 0.0010916125029325485
step: 130, loss: 0.001586219179444015
step: 140, loss: 0.0022148967254906893
step: 150, loss: 0.0006220002542249858
step: 160, loss: 0.0010194857604801655
step: 170, loss: 0.0011923181591555476
step: 180, loss: 0.010465383529663086
step: 190, loss: 0.0001632122293813154
step: 200, loss: 0.0013701559510082006
step: 210, loss: 0.003933302592486143
step: 220, loss: 0.0015761548420414329
step: 230, loss: 0.00032666471088305116
step: 240, loss: 0.0005168510251678526
step: 250, loss: 0.0036687033716589212
step: 260, loss: 0.0001979904918698594
step: 270, loss: 0.033314693719148636
step: 280, loss: 0.00821802020072937
step: 290, loss: 0.0011843318352475762
step: 300, loss: 4.9583526561036706e-05
step: 310, loss: 0.014687690883874893
step: 320, loss: 0.0008058559033088386
step: 330, loss: 0.0009686945704743266
step: 340, loss: 0.004434369970113039
step: 350, loss: 0.03342191129922867
step: 360, loss: 0.001648598350584507
step: 370, loss: 0.0035744053311645985
step: 380, loss: 0.0046987892128527164
step: 390, loss: 0.00016570401203352958
step: 400, loss: 0.01804722473025322
```

```

step: 410, loss: 0.005330492742359638
step: 420, loss: 0.007219989784061909
step: 430, loss: 0.003895118599757552
step: 440, loss: 0.001074014464393258
step: 450, loss: 0.01742270216345787
step: 460, loss: 0.0003272337489761412
step: 470, loss: 0.019054502248764038
step: 480, loss: 8.236078429035842e-05
step: 490, loss: 0.008235668763518333
step: 500, loss: 0.0009524050401523709
step: 510, loss: 0.09146641194820404
step: 520, loss: 0.002571417484432459
step: 530, loss: 0.07977344840765
/usr/local/lib/python3.7/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=0.9464368886818817, f1=0.9387186629526463,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011019648518413305
step: 10, loss: 0.0011096912203356624
step: 20, loss: 0.00037904366035945714
step: 30, loss: 0.003543023020029068
step: 40, loss: 0.0038753810804337263
step: 50, loss: 0.0011779807973653078
step: 60, loss: 0.017380928620696068
step: 70, loss: 0.007621806114912033
step: 80, loss: 0.001387071912176907
step: 90, loss: 0.006735313683748245
step: 100, loss: 0.0007744583999738097
step: 110, loss: 0.009941195137798786
step: 120, loss: 7.709467899985611e-05
step: 130, loss: 3.6002547858515754e-05
step: 140, loss: 0.00034708782914094627
step: 150, loss: 0.0016489882254973054
step: 160, loss: 0.028180843219161034
step: 170, loss: 0.00011254784476477653
step: 180, loss: 0.0967201218008995
step: 190, loss: 3.174308221787214e-05
step: 200, loss: 0.0003340561524964869
step: 210, loss: 0.013782596215605736
step: 220, loss: 0.0005492665222845972
step: 230, loss: 0.0010388484224677086
step: 240, loss: 0.0012008694466203451
step: 250, loss: 0.0063714757561683655
step: 260, loss: 0.009553574956953526

```

```

step: 270, loss: 0.005015430506318808
step: 280, loss: 0.027589788660407066
step: 290, loss: 0.01998789794743061
step: 300, loss: 0.04831964150071144
step: 310, loss: 0.017288682982325554
step: 320, loss: 0.0288830678910017
step: 330, loss: 0.004391706082969904
step: 340, loss: 9.950394451152533e-05
step: 350, loss: 0.000586509529966861
step: 360, loss: 0.005106027238070965
step: 370, loss: 0.0033167111687362194
step: 380, loss: 0.015364197082817554
step: 390, loss: 0.000987196690402925
step: 400, loss: 0.00047461941721849144
step: 410, loss: 0.0016917112516239285
step: 420, loss: 0.0006501654861494899
step: 430, loss: 0.0010394565761089325
step: 440, loss: 0.0005033882334828377
step: 450, loss: 0.002725919708609581
step: 460, loss: 0.00021603572531603277
step: 470, loss: 0.002581023843958974
step: 480, loss: 0.0017012248281389475
step: 490, loss: 0.00106367701664567
step: 500, loss: 0.0045937844552099705
step: 510, loss: 0.0012043163878843188
step: 520, loss: 0.0020562249701470137
step: 530, loss: 0.01772071234881878
/usr/local/lib/python3.7/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=0.9441860465116279, f1=0.9456572224802601,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010393512202426791
step: 10, loss: 0.0008013302576728165
step: 20, loss: 0.001594889909029007
step: 30, loss: 0.0012602830538526177
step: 40, loss: 0.000827243085950613
step: 50, loss: 0.00046439748257398605
step: 60, loss: 0.0008816817426122725
step: 70, loss: 0.0003803735598921776
step: 80, loss: 0.0008017220534384251
step: 90, loss: 0.02824646793305874
step: 100, loss: 0.0009416568209417164
step: 110, loss: 0.030266549438238144
step: 120, loss: 0.05125528201460838

```



```
step: 130, loss: 0.0003803988511208445
step: 140, loss: 0.0006547037046402693
step: 150, loss: 0.014969924464821815
step: 160, loss: 0.007360141724348068
step: 170, loss: 0.0016672657802700996
step: 180, loss: 5.238961966824718e-05
step: 190, loss: 0.0005137899424880743
step: 200, loss: 0.0005041087861172855
step: 210, loss: 0.001000786549411714
step: 220, loss: 0.05376293137669563
step: 230, loss: 0.0028927272651344538
step: 240, loss: 0.0021266506519168615
step: 250, loss: 2.2007958250469528e-05
step: 260, loss: 0.004342691507190466
step: 270, loss: 0.001575217698700726
step: 280, loss: 0.0016079460037872195
step: 290, loss: 0.0381520576775074
step: 300, loss: 0.00024461280554533005
step: 310, loss: 0.026430681347846985
step: 320, loss: 0.004552175290882587
step: 330, loss: 0.0012475945986807346
step: 340, loss: 0.005385211668908596
step: 350, loss: 0.0003407453768886626
step: 360, loss: 0.0008931186748668551
step: 370, loss: 0.0028606040868908167
step: 380, loss: 0.0011067073792219162
step: 390, loss: 0.0016131665324792266
step: 400, loss: 0.00012824639270547777
step: 410, loss: 0.000956783420406282
step: 420, loss: 0.01971958950161934
step: 430, loss: 0.0006491238600574434
step: 440, loss: 0.00039926619501784444
step: 450, loss: 0.0024630350526422262
step: 460, loss: 0.0033210469409823418
step: 470, loss: 0.0008854490588419139
step: 480, loss: 0.00012481991143431515
step: 490, loss: 0.0017012576572597027
step: 500, loss: 0.00020542244601529092
step: 510, loss: 0.0038531741593033075
step: 520, loss: 0.00017180187569465488
step: 530, loss: 0.010655799880623817
/usr/local/lib/python3.7/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=0.9386814200092208, f1=0.9407441433164906,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005416920757852495
step: 10, loss: 0.0030191789846867323
step: 20, loss: 0.008045295253396034
step: 30, loss: 0.0001804508501663804
step: 40, loss: 0.021252600476145744
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.005906923674046993
step: 60, loss: 5.5472231906605884e-05
step: 70, loss: 0.08517491817474365
step: 80, loss: 0.0006654148455709219
step: 90, loss: 0.002410299377515912
step: 100, loss: 0.048734769225120544
step: 110, loss: 0.01803717389702797
step: 120, loss: 0.004966528620570898
step: 130, loss: 0.02160760946571827
step: 140, loss: 0.0004012945864815265
step: 150, loss: 0.009529437869787216
step: 160, loss: 0.011669199913740158
step: 170, loss: 1.084052291844273e-05
step: 180, loss: 0.00018498986901249737
step: 190, loss: 0.0009715129854157567
step: 200, loss: 0.003256396623328328
step: 210, loss: 2.079015212075319e-05
step: 220, loss: 0.00020653038518503308
step: 230, loss: 9.366277663502842e-05
step: 240, loss: 0.0002324259257875383
step: 250, loss: 9.87565545074176e-06
step: 260, loss: 0.0002930421906057745
step: 270, loss: 0.05288795381784439
step: 280, loss: 0.0001231919159181416
step: 290, loss: 8.935870573623106e-05
step: 300, loss: 0.0019279164262115955
step: 310, loss: 0.0007368004880845547
step: 320, loss: 0.007244590204209089
step: 330, loss: 0.04057278856635094
step: 340, loss: 0.0004211770719848573
step: 350, loss: 8.124815394694451e-06
step: 360, loss: 3.395224121049978e-05
step: 370, loss: 0.00016258655523415655
step: 380, loss: 0.0009590075351297855
step: 390, loss: 0.05646579712629318
step: 400, loss: 0.0001541431265650317
step: 410, loss: 0.010519666597247124
step: 420, loss: 0.14760597050189972
step: 430, loss: 0.008799615316092968
step: 440, loss: 0.08824067562818527
step: 450, loss: 0.0003940226451959461
```

```

step: 460, loss: 4.3568819819483906e-05
step: 470, loss: 0.0021325796842575073
step: 480, loss: 0.025933345779776573
step: 490, loss: 0.0011634171241894364
step: 500, loss: 0.001192216295748949
step: 510, loss: 0.02350088767707348
step: 520, loss: 0.0020621041767299175
step: 530, loss: 0.0015763159608468413
/usr/local/lib/python3.7/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=0.9460966542750929, f1=0.9435185185185185,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001718884304864332
step: 10, loss: 0.0008498140960000455
step: 20, loss: 0.0012137622106820345
step: 30, loss: 0.0004610228934325278
step: 40, loss: 0.005300554446876049
step: 50, loss: 0.0035548065789043903
step: 60, loss: 0.0042627532966434956
step: 70, loss: 0.0026377784088253975
step: 80, loss: 0.004655235446989536
step: 90, loss: 0.004406034015119076
step: 100, loss: 0.0003870803629979491
step: 110, loss: 0.0007479244377464056
step: 120, loss: 5.507865716936067e-05
step: 130, loss: 0.0025708647444844246
step: 140, loss: 0.0012015984393656254
step: 150, loss: 0.0025784350000321865
step: 160, loss: 0.0003541026380844414
step: 170, loss: 0.006263740826398134
step: 180, loss: 0.00042538391426205635
step: 190, loss: 0.00013398431474342942
step: 200, loss: 0.00024383066920563579
step: 210, loss: 3.139975160593167e-05
step: 220, loss: 0.024809470400214195
step: 230, loss: 0.0035514351911842823
step: 240, loss: 0.04478289932012558
step: 250, loss: 0.0007616423536092043
step: 260, loss: 0.0016656778752803802
step: 270, loss: 0.0008718307944945991
step: 280, loss: 6.233088788576424e-05
step: 290, loss: 0.0009173667640425265
step: 300, loss: 0.0001924815442180261
step: 310, loss: 0.00041824832442216575

```

```

step: 320, loss: 0.00018929000361822546
step: 330, loss: 0.0015671945875510573
step: 340, loss: 0.0018706333357840776
step: 350, loss: 0.0009073209948837757
step: 360, loss: 0.09006711840629578
step: 370, loss: 0.009352491237223148
step: 380, loss: 0.0010864357464015484
step: 390, loss: 0.0007525191176682711
step: 400, loss: 0.0010410556569695473
step: 410, loss: 0.0009712653118185699
step: 420, loss: 0.0003228779823984951
step: 430, loss: 6.002683949191123e-05
step: 440, loss: 0.000119554199045524
step: 450, loss: 0.00022963466471992433
step: 460, loss: 0.0002694023714866489
step: 470, loss: 0.0020900273229926825
step: 480, loss: 3.356425077072345e-05
step: 490, loss: 0.00011003702820744365
step: 500, loss: 0.00414184108376503
step: 510, loss: 8.713373972568661e-06
step: 520, loss: 0.0001759072911227122
step: 530, loss: 2.7810481697088107e-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=0.9484440315838365, f1=0.9425393883225208,
best_f1=0.9399720800372265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033210739493370056
step: 10, loss: 0.0006740076933056116
step: 20, loss: 0.00012021810107398778
step: 30, loss: 0.00026716195861808956
step: 40, loss: 0.00010406568617327139
step: 50, loss: 0.001000580727122724
step: 60, loss: 0.014827493578195572
step: 70, loss: 0.002666595159098506
step: 80, loss: 0.0001376466389046982
step: 90, loss: 6.31857619737275e-05
step: 100, loss: 8.962559513747692e-05
step: 110, loss: 0.00041654842789284885
step: 120, loss: 1.003570287139155e-05
step: 130, loss: 1.5652429283363745e-05
step: 140, loss: 0.008842481300234795
step: 150, loss: 0.002394341630861163
step: 160, loss: 0.0023133072536438704
step: 170, loss: 0.001966077834367752

```

```

step: 180, loss: 0.00022811314556747675
step: 190, loss: 0.00267805065959692
step: 200, loss: 0.0008573210216127336
step: 210, loss: 0.008236976340413094
step: 220, loss: 1.3593332369055133e-05
step: 230, loss: 0.021153178066015244
step: 240, loss: 0.0027509192004799843
step: 250, loss: 0.0013288843911141157
step: 260, loss: 0.00025193122564814985
step: 270, loss: 0.0006665404653176665
step: 280, loss: 0.001106725656427443
step: 290, loss: 8.139306009979919e-05
step: 300, loss: 2.205930468335282e-05
step: 310, loss: 0.0001768874062690884
step: 320, loss: 2.237163243989926e-05
step: 330, loss: 0.001477074809372425
step: 340, loss: 0.0004230674239806831
step: 350, loss: 0.0005607568309642375
step: 360, loss: 0.0017032341565936804
step: 370, loss: 0.0002616470737848431
step: 380, loss: 0.026149094104766846
step: 390, loss: 0.0011412266176193953
step: 400, loss: 0.0019973262678831816
step: 410, loss: 9.782781853573397e-05
step: 420, loss: 0.00020019781368318945
step: 430, loss: 0.0008946319576352835
step: 440, loss: 0.0010601809481158853
step: 450, loss: 0.00021110981469973922
step: 460, loss: 0.048686861991882324
step: 470, loss: 2.8937356546521187e-05
step: 480, loss: 0.00035304256016388535
step: 490, loss: 0.0003169897827319801
step: 500, loss: 0.0022023681085556746
step: 510, loss: 0.00010166376887355
step: 520, loss: 0.0008445028797723353
step: 530, loss: 0.0017690520035102963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9488372093023255, f1=0.9427640763145649,
best_f1=0.9427640763145649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 2.3740662072668783e-05
step: 10, loss: 0.001756120822392404
step: 20, loss: 0.0032718167640268803
step: 30, loss: 0.005966573487967253
step: 40, loss: 0.0020950925536453724
step: 50, loss: 0.0010154034243896604
step: 60, loss: 0.0005187590140849352
step: 70, loss: 0.00010132090392289683
step: 80, loss: 0.00021035855752415955
step: 90, loss: 9.197795588988811e-05
step: 100, loss: 4.314568650443107e-05
step: 110, loss: 0.004734191577881575
step: 120, loss: 0.00032148530590347946
step: 130, loss: 0.0011030384339392185
step: 140, loss: 2.4742164896451868e-05
step: 150, loss: 0.0018957266584038734
step: 160, loss: 0.0002765564131550491
step: 170, loss: 3.582668796298094e-05
step: 180, loss: 0.0004650465853046626
step: 190, loss: 0.019813258200883865
step: 200, loss: 0.0018813046626746655
step: 210, loss: 0.0009020080324262381
step: 220, loss: 0.002054166281595826
step: 230, loss: 0.0003474121622275561
step: 240, loss: 0.00037259457167237997
step: 250, loss: 0.0007393735577352345
step: 260, loss: 1.848818465077784e-05
step: 270, loss: 1.8200702470494434e-05
step: 280, loss: 0.0003641430812422186
step: 290, loss: 0.0002904823049902916
step: 300, loss: 0.0003690998419187963
step: 310, loss: 0.033612485975027084
step: 320, loss: 0.00870937667787075
step: 330, loss: 1.620839975657873e-05
step: 340, loss: 0.0001594827335793525
step: 350, loss: 0.007523437961935997
step: 360, loss: 0.00035103957634419203
step: 370, loss: 0.0014760716585442424
step: 380, loss: 0.0016408697701990604
step: 390, loss: 0.0002358187921345234
step: 400, loss: 0.046700023114681244
step: 410, loss: 0.00014051454490981996
step: 420, loss: 0.00023291638353839517
step: 430, loss: 1.31240831251489e-05
step: 440, loss: 0.002349437214434147
step: 450, loss: 0.0018351614708080888
step: 460, loss: 0.0031108027324080467
step: 470, loss: 3.680684312712401e-05

```

step: 480, loss: 0.0028766789473593235
step: 490, loss: 0.0007081442163325846
step: 500, loss: 0.0028623396065086126
step: 510, loss: 0.0005862120306119323
step: 520, loss: 5.3974876209395006e-05
step: 530, loss: 0.0005201288149692118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9492314857941314, f1=0.9438515081206497,
best_f1=0.9438515081206497
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 [00:34, 168.61it/s]
load_f1 = 0.950420954162769
real_f1 = 0.9510032664489033

```

```
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, 140.21it/s]
```

4.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:
 enabled : True


```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5011301040649414
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.392741322517395
step: 20, loss: 0.44149646162986755
step: 30, loss: 0.28895673155784607
step: 40, loss: 0.3257209360599518
step: 50, loss: 0.4405916929244995
step: 60, loss: 0.5206957459449768
step: 70, loss: 0.30991455912590027
step: 80, loss: 0.3594716191291809
step: 90, loss: 0.20611999928951263
step: 100, loss: 0.26736995577812195
step: 110, loss: 0.28333747386932373
step: 120, loss: 0.3862968385219574
step: 130, loss: 0.2909175455570221
step: 140, loss: 0.4689903259277344
step: 150, loss: 0.3101464509963989

```

```

step: 160, loss: 0.5036044120788574
step: 170, loss: 0.2471337765455246
step: 180, loss: 0.4061252176761627
step: 190, loss: 0.6458840370178223
step: 200, loss: 0.35336530208587646
step: 210, loss: 0.4897250831127167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3947436511516571
step: 10, loss: 0.19612517952919006
step: 20, loss: 0.5242892503738403
step: 30, loss: 0.49789705872535706
step: 40, loss: 0.45098644495010376
step: 50, loss: 0.23830491304397583
step: 60, loss: 0.30653733015060425
step: 70, loss: 0.4419209659099579
step: 80, loss: 0.31485021114349365
step: 90, loss: 0.377852201461792
step: 100, loss: 0.4885789155960083
step: 110, loss: 0.3902382552623749
step: 120, loss: 0.23774540424346924
step: 130, loss: 0.17385733127593994
step: 140, loss: 0.25331830978393555
step: 150, loss: 0.4273577332496643
step: 160, loss: 0.17344622313976288
step: 170, loss: 0.5708234310150146
step: 180, loss: 0.3014583885669708
step: 190, loss: 0.2965922951698303
step: 200, loss: 0.14939595758914948
step: 210, loss: 0.3289080262184143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2167729279058362, f1=0.21613691931540344,

```

```

best_f1=0.21613691931540344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23383620381355286
step: 10, loss: 0.2261795550584793
step: 20, loss: 0.4571560025215149
step: 30, loss: 0.25982901453971863
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.5110940337181091
step: 50, loss: 0.43725740909576416
step: 60, loss: 0.507220983505249
step: 70, loss: 0.2237580567598343
step: 80, loss: 0.45031020045280457
step: 90, loss: 0.24039466679096222
step: 100, loss: 0.39209914207458496
step: 110, loss: 0.23125125467777252
step: 120, loss: 0.2203129082918167
step: 130, loss: 0.21664494276046753
step: 140, loss: 0.3610980808734894
step: 150, loss: 0.25611013174057007
step: 160, loss: 0.22505584359169006
step: 170, loss: 0.35321617126464844
step: 180, loss: 0.29992586374282837
step: 190, loss: 0.1664671152830124
step: 200, loss: 0.22435739636421204
step: 210, loss: 0.2802746891975403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2323232323232323, f1=0.22535211267605634,
best_f1=0.22535211267605634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3234795331954956
step: 10, loss: 0.283830851316452
step: 20, loss: 0.27005016803741455
step: 30, loss: 0.2255268096923828
step: 40, loss: 0.19011689722537994
step: 50, loss: 0.2235378473997116
step: 60, loss: 0.4585704207420349
step: 70, loss: 0.24064795672893524
step: 80, loss: 0.23027303814888

```

```

step: 90, loss: 0.3767027258872986
step: 100, loss: 0.3445942997932434
step: 110, loss: 0.6447015404701233
step: 120, loss: 0.33716729283332825
step: 130, loss: 0.6761392951011658
step: 140, loss: 0.43267232179641724
step: 150, loss: 0.3459579050540924
step: 160, loss: 0.31400617957115173
step: 170, loss: 0.21880091726779938
step: 180, loss: 0.07834969460964203
step: 190, loss: 0.16618020832538605
step: 200, loss: 0.29763850569725037
step: 210, loss: 0.44044047594070435
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2531089978054133, f1=0.262992125984252,
best_f1=0.262992125984252
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33618083596229553
step: 10, loss: 0.33296382427215576
step: 20, loss: 0.30615881085395813
step: 30, loss: 0.2367420494556427
step: 40, loss: 0.39903101325035095
step: 50, loss: 0.36926132440567017
step: 60, loss: 0.38533860445022583
step: 70, loss: 0.25451451539993286
step: 80, loss: 0.41172224283218384
step: 90, loss: 0.37729862332344055
step: 100, loss: 0.2012505829334259
step: 110, loss: 0.1621803194284439
step: 120, loss: 0.2631208598613739
step: 130, loss: 0.2662941813468933
step: 140, loss: 0.48208001255989075
step: 150, loss: 0.2285022735595703
step: 160, loss: 0.19549071788787842
step: 170, loss: 0.3324127197265625
step: 180, loss: 0.1961325854063034
step: 190, loss: 0.3664124310016632
step: 200, loss: 0.4489586055278778
step: 210, loss: 0.17064017057418823
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.38408304498269896, f1=0.34307992202729043,
best_f1=0.34307992202729043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08751457929611206
step: 10, loss: 0.3888103663921356
step: 20, loss: 0.25144854187965393
step: 30, loss: 0.26586002111434937
step: 40, loss: 0.2622690200805664
step: 50, loss: 0.5548990368843079
step: 60, loss: 0.2613653242588043
step: 70, loss: 0.3456341624259949
step: 80, loss: 0.2763305902481079
step: 90, loss: 0.33100780844688416
step: 100, loss: 0.35170403122901917
step: 110, loss: 0.3150678873062134
step: 120, loss: 0.2049165517091751
step: 130, loss: 0.1745574027299881
step: 140, loss: 0.3685385286808014
step: 150, loss: 0.1603647768497467
step: 160, loss: 0.11768542975187302
step: 170, loss: 0.4022179841995239
step: 180, loss: 0.30048441886901855
step: 190, loss: 0.23590746521949768
step: 200, loss: 0.35030457377433777
step: 210, loss: 0.2476889193058014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4045676998368678, f1=0.39542483660130723,
best_f1=0.39542483660130723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3688492774963379
step: 10, loss: 0.19004447758197784
step: 20, loss: 0.41049832105636597

```

```

step: 30, loss: 0.18954139947891235
step: 40, loss: 0.2033160775899887
step: 50, loss: 0.2901553809642792
step: 60, loss: 0.2771437466144562
step: 70, loss: 0.1963931769132614
step: 80, loss: 0.34725186228752136
step: 90, loss: 0.3558911085128784
step: 100, loss: 0.27116233110427856
step: 110, loss: 0.22938191890716553
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 120, loss: 0.2751947045326233
step: 130, loss: 0.38844600319862366
step: 140, loss: 0.2797927260398865
step: 150, loss: 0.29025998711586
step: 160, loss: 0.615401029586792
step: 170, loss: 0.45735061168670654
step: 180, loss: 0.20187394320964813
step: 190, loss: 0.2497951090335846
step: 200, loss: 0.24435852468013763
step: 210, loss: 0.44579610228538513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45029239766081863, f1=0.45538461538461533,
best_f1=0.45538461538461533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37403199076652527
step: 10, loss: 0.2405380755662918
step: 20, loss: 0.15845024585723877
step: 30, loss: 0.20497210323810577
step: 40, loss: 0.28127503395080566
step: 50, loss: 0.07948990166187286
step: 60, loss: 0.09182386100292206
step: 70, loss: 0.321272075176239
step: 80, loss: 0.2410350739955902
step: 90, loss: 0.35967501997947693
step: 100, loss: 0.3704088628292084
step: 110, loss: 0.24347876012325287
step: 120, loss: 0.23214401304721832
step: 130, loss: 0.05979140102863312
step: 140, loss: 0.23603834211826324
step: 150, loss: 0.35098740458488464

```

```

step: 160, loss: 0.44349634647369385
step: 170, loss: 0.43581220507621765
step: 180, loss: 0.2000882774591446
step: 190, loss: 0.11687412858009338
step: 200, loss: 0.20685356855392456
step: 210, loss: 0.3646913468837738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.466887417218543, f1=0.508361204013378,
best_f1=0.508361204013378
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36883339285850525
step: 10, loss: 0.25110599398612976
step: 20, loss: 0.29223254323005676
step: 30, loss: 0.07976295799016953
step: 40, loss: 0.09496256709098816
step: 50, loss: 0.4225676655769348
step: 60, loss: 0.09428815543651581
step: 70, loss: 0.17140506207942963
step: 80, loss: 0.10588499903678894
step: 90, loss: 0.2696187496185303
step: 100, loss: 0.2812052071094513
step: 110, loss: 0.2438315898180008
step: 120, loss: 0.31436794996261597
step: 130, loss: 0.1373656988143921
step: 140, loss: 0.3177257180213928
step: 150, loss: 0.0956415981054306
step: 160, loss: 0.2618246376514435
step: 170, loss: 0.20299848914146423
step: 180, loss: 0.3024667203426361
step: 190, loss: 0.11391552537679672
step: 200, loss: 0.14171497523784637
step: 210, loss: 0.2173279970884323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5719237435008666, f1=0.6181172291296626,

```

```

best_f1=0.6181172291296626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13786953687667847
step: 10, loss: 0.2045954316854477
step: 20, loss: 0.11155945062637329
step: 30, loss: 0.08461833000183105
step: 40, loss: 0.15843479335308075
step: 50, loss: 0.19439958035945892
step: 60, loss: 0.1139703020453453
step: 70, loss: 0.16714514791965485
step: 80, loss: 0.16494764387607574
step: 90, loss: 0.15443657338619232
step: 100, loss: 0.32333070039749146
step: 110, loss: 0.05766941234469414
step: 120, loss: 0.462240993976593
step: 130, loss: 0.09302252531051636
step: 140, loss: 0.14762389659881592
step: 150, loss: 0.15600039064884186
step: 160, loss: 0.08614084869623184
step: 170, loss: 0.05591551959514618
step: 180, loss: 0.13042345643043518
step: 190, loss: 0.11125023663043976
step: 200, loss: 0.3179999589920044
step: 210, loss: 0.1456880420446396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5862785862785863, f1=0.5882352941176471,
best_f1=0.5882352941176471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10467777401208878
step: 10, loss: 0.18448664247989655
step: 20, loss: 0.0687481015920639
step: 30, loss: 0.13422542810440063
step: 40, loss: 0.1814388930797577
step: 50, loss: 0.37483009696006775
step: 60, loss: 0.16461892426013947
step: 70, loss: 0.10320284962654114
step: 80, loss: 0.2301318198442459
step: 90, loss: 0.239568293094635

```



```

step: 100, loss: 0.19552625715732574
step: 110, loss: 0.39231643080711365
step: 120, loss: 0.2061842679977417
step: 130, loss: 0.09703681617975235
step: 140, loss: 0.1345282346010208
step: 150, loss: 0.17693912982940674
step: 160, loss: 0.03431306034326553
step: 170, loss: 0.14454330503940582
step: 180, loss: 0.1649475246667862
step: 190, loss: 0.2987404465675354
step: 200, loss: 0.06870967894792557
step: 210, loss: 0.1805436760187149
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5870841487279843, f1=0.6015037593984963,
best_f1=0.6015037593984963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09990900754928589
step: 10, loss: 0.14214661717414856
step: 20, loss: 0.2508689761161804
step: 30, loss: 0.10374996811151505
step: 40, loss: 0.06879756599664688
step: 50, loss: 0.1834484487771988
step: 60, loss: 0.09959220886230469
step: 70, loss: 0.22803714871406555
step: 80, loss: 0.11473318934440613
step: 90, loss: 0.26551735401153564
step: 100, loss: 0.038478605449199677
step: 110, loss: 0.18449756503105164
step: 120, loss: 0.037998978048563004
step: 130, loss: 0.13889454305171967
step: 140, loss: 0.3440394997596741
step: 150, loss: 0.02511814795434475
step: 160, loss: 0.10579463839530945
step: 170, loss: 0.23210792243480682
step: 180, loss: 0.10918058454990387
step: 190, loss: 0.11919084936380386
step: 200, loss: 0.05314529314637184
step: 210, loss: 0.08226174116134644
/usr/local/lib/python3.7/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=0.5852713178294574, f1=0.6132812500000001,
best_f1=0.6015037593984963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08774582296609879
step: 10, loss: 0.03120950050652027
step: 20, loss: 0.06438732892274857
step: 30, loss: 0.07687344402074814
step: 40, loss: 0.1701233983039856
step: 50, loss: 0.1011376827955246
step: 60, loss: 0.22070841491222382
step: 70, loss: 0.0580952949821949
step: 80, loss: 0.21306493878364563
step: 90, loss: 0.22786280512809753
step: 100, loss: 0.11100469529628754
step: 110, loss: 0.1383911371231079
step: 120, loss: 0.09449122101068497
step: 130, loss: 0.0392831452190876
step: 140, loss: 0.1503104716539383
step: 150, loss: 0.14400233328342438
step: 160, loss: 0.12205982953310013
step: 170, loss: 0.0670977458357811
step: 180, loss: 0.035241831094026566
step: 190, loss: 0.07642853260040283
step: 200, loss: 0.0930481106042862
step: 210, loss: 0.12365309149026871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.600375234521576, f1=0.6303939962476548,
best_f1=0.6303939962476548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01580742560327053
step: 10, loss: 0.06905801594257355
step: 20, loss: 0.16898073256015778
step: 30, loss: 0.02300153486430645
step: 40, loss: 0.12355843186378479
step: 50, loss: 0.07078752666711807
step: 60, loss: 0.19675908982753754
step: 70, loss: 0.13316892087459564

```

```

step: 80, loss: 0.09116436541080475
step: 90, loss: 0.03927777707576752
step: 100, loss: 0.03548962622880936
step: 110, loss: 0.17129945755004883
step: 120, loss: 0.03974289819598198
step: 130, loss: 0.05046622455120087
step: 140, loss: 0.11438176035881042
step: 150, loss: 0.13149720430374146
step: 160, loss: 0.09750247001647949
step: 170, loss: 0.0924677774310112
step: 180, loss: 0.05589789152145386
step: 190, loss: 0.13836334645748138
step: 200, loss: 0.05225968733429909
step: 210, loss: 0.03476965054869652
/usr/local/lib/python3.7/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=0.5967078189300411, f1=0.6237424547283703,
best_f1=0.6303939962476548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07128214836120605
step: 10, loss: 0.08281781524419785
step: 20, loss: 0.17198599874973297
step: 30, loss: 0.1060311421751976
step: 40, loss: 0.20753273367881775
step: 50, loss: 0.12983380258083344
step: 60, loss: 0.21059304475784302
step: 70, loss: 0.08508757501840591
step: 80, loss: 0.27973201870918274
step: 90, loss: 0.05300929397344589
step: 100, loss: 0.08735425770282745
step: 110, loss: 0.11046550422906876
step: 120, loss: 0.2321690171957016
step: 130, loss: 0.10764671117067337
step: 140, loss: 0.09022337198257446
step: 150, loss: 0.25942301750183105
step: 160, loss: 0.10832173377275467
step: 170, loss: 0.11957346647977829
step: 180, loss: 0.07166190445423126
step: 190, loss: 0.19786356389522552
step: 200, loss: 0.026024116203188896
step: 210, loss: 0.24228531122207642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.6065573770491803, f1=0.6372745490981965,
best_f1=0.6372745490981965
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, 231.59it/s]
load_f1 = 0.5735294117647058
real_f1 = 0.5650557620817844
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.16it/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: 0.46744653582572937
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41393136978149414
step: 20, loss: 0.25977447628974915
step: 30, loss: 0.3676232397556305
step: 40, loss: 0.24466176331043243
step: 50, loss: 0.30821937322616577
step: 60, loss: 0.4532616138458252
step: 70, loss: 0.47150784730911255
step: 80, loss: 0.13873308897018433
step: 90, loss: 0.31394392251968384
step: 100, loss: 0.4249928295612335
step: 110, loss: 0.23392236232757568
step: 120, loss: 0.2964828312397003
step: 130, loss: 0.33533498644828796
step: 140, loss: 0.17481926083564758
step: 150, loss: 0.3305782675743103
step: 160, loss: 0.2325257658958435
step: 170, loss: 0.37359631061553955
step: 180, loss: 0.17324253916740417
step: 190, loss: 0.1706366240978241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.38239696621894836
step: 10, loss: 0.31537097692489624
step: 20, loss: 0.6689420342445374
step: 30, loss: 0.2597464323043823
step: 40, loss: 0.5669799447059631
step: 50, loss: 0.32314878702163696
step: 60, loss: 0.4528667628765106
step: 70, loss: 0.31035879254341125
step: 80, loss: 0.17141291499137878
step: 90, loss: 0.30480527877807617
step: 100, loss: 0.24539199471473694
step: 110, loss: 0.3952009379863739
step: 120, loss: 0.23327261209487915
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5534840822219849
step: 140, loss: 0.31851014494895935
step: 150, loss: 0.3012409508228302
step: 160, loss: 0.3137798607349396
step: 170, loss: 0.2438904345035553
step: 180, loss: 0.17774373292922974
step: 190, loss: 0.23183034360408783
/usr/local/lib/python3.7/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=0.17216770740410348, 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: 0.38025811314582825
step: 10, loss: 0.3890708088874817
step: 20, loss: 0.4776931405067444
step: 30, loss: 0.3104410469532013
step: 40, loss: 0.08333691209554672
step: 50, loss: 0.37867921590805054
step: 60, loss: 0.16261635720729828
step: 70, loss: 0.3801961839199066
step: 80, loss: 0.3306877017021179

```

```

step: 90, loss: 0.3767678439617157
step: 100, loss: 0.527411699295044
step: 110, loss: 0.6669667959213257
step: 120, loss: 0.3767663538455963
step: 130, loss: 0.15590117871761322
step: 140, loss: 0.37651512026786804
step: 150, loss: 0.31058168411254883
step: 160, loss: 0.6298076510429382
step: 170, loss: 0.44800782203674316
step: 180, loss: 0.37996795773506165
step: 190, loss: 0.16028639674186707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.23376372456550598
step: 10, loss: 0.23352785408496857
step: 20, loss: 0.31099066138267517
step: 30, loss: 0.24782051146030426
step: 40, loss: 0.5425788760185242
step: 50, loss: 0.24466556310653687
step: 60, loss: 0.37936243414878845
step: 70, loss: 0.31194254755973816
step: 80, loss: 0.24940112233161926
step: 90, loss: 0.17628227174282074
step: 100, loss: 0.31003254652023315
step: 110, loss: 0.3755688965320587
step: 120, loss: 0.2490304559469223
step: 130, loss: 0.45998165011405945
step: 140, loss: 0.3106916844844818
step: 150, loss: 0.24649158120155334
step: 160, loss: 0.30391302704811096
step: 170, loss: 0.4316120743751526
step: 180, loss: 0.3806852102279663
step: 190, loss: 0.1607212871313095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/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: 0.4027289152145386

```



```

step: 10, loss: 0.36649665236473083
step: 20, loss: 0.1578369289636612
step: 30, loss: 0.1199699267745018
step: 40, loss: 0.30706366896629333
step: 50, loss: 0.5077413320541382
step: 60, loss: 0.23619024455547333
step: 70, loss: 0.39276638627052307
step: 80, loss: 0.34698718786239624
step: 90, loss: 0.2742856740951538
step: 100, loss: 0.444031685590744
step: 110, loss: 0.40527695417404175
step: 120, loss: 0.24129436910152435
step: 130, loss: 0.6094657182693481
step: 140, loss: 0.3455657660961151
step: 150, loss: 0.2717099189758301
step: 160, loss: 0.15522736310958862
step: 170, loss: 0.37911415100097656
step: 180, loss: 0.23969294130802155
step: 190, loss: 0.2879285514354706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2017353579175705, f1=0.20501635768811338,
best_f1=0.20501635768811338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2872099280357361
step: 10, loss: 0.23756049573421478
step: 20, loss: 0.2761126458644867
step: 30, loss: 0.4468994140625
step: 40, loss: 0.2827906310558319
step: 50, loss: 0.29179349541664124
step: 60, loss: 0.42119839787483215
step: 70, loss: 0.29643967747688293
step: 80, loss: 0.29227694869041443
step: 90, loss: 0.21512582898139954
step: 100, loss: 0.45278191566467285
step: 110, loss: 0.22981467843055725
step: 120, loss: 0.4254300594329834
step: 130, loss: 0.5757023692131042
step: 140, loss: 0.17315351963043213
step: 150, loss: 0.3722694218158722
step: 160, loss: 0.37452346086502075

```

```

step: 170, loss: 0.36564064025878906
step: 180, loss: 0.17888039350509644
step: 190, loss: 0.31043198704719543
/usr/local/lib/python3.7/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=0.2017353579175705, f1=0.20501635768811338,
best_f1=0.20501635768811338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30092722177505493
step: 10, loss: 0.2813752293586731
step: 20, loss: 0.2788507342338562
step: 30, loss: 0.20131905376911163
step: 40, loss: 0.2905016541481018
step: 50, loss: 0.08591541647911072
step: 60, loss: 0.1461159735918045
step: 70, loss: 0.14952099323272705
step: 80, loss: 0.23136359453201294
step: 90, loss: 0.23028238117694855
step: 100, loss: 0.5298022627830505
step: 110, loss: 0.4274841845035553
step: 120, loss: 0.35672444105148315
step: 130, loss: 0.28357505798339844
step: 140, loss: 0.2195817232131958
step: 150, loss: 0.27832096815109253
step: 160, loss: 0.3650111258029938
step: 170, loss: 0.28259459137916565
step: 180, loss: 0.21702182292938232
step: 190, loss: 0.3105620741844177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2029621503017005, f1=0.2056384742951907,
best_f1=0.2056384742951907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22685867547988892
step: 10, loss: 0.3555240333080292
step: 20, loss: 0.221824049949646
step: 30, loss: 0.36390185356140137
step: 40, loss: 0.15906177461147308

```

```

step: 50, loss: 0.29196685552597046
step: 60, loss: 0.42839837074279785
step: 70, loss: 0.21633818745613098
step: 80, loss: 0.4009239673614502
step: 90, loss: 0.2200249582529068
step: 100, loss: 0.2693406343460083
step: 110, loss: 0.36315906047821045
step: 120, loss: 0.36866435408592224
step: 130, loss: 0.3012605905532837
step: 140, loss: 0.33360370993614197
step: 150, loss: 0.2926342487335205
step: 160, loss: 0.17460954189300537
step: 170, loss: 0.22806166112422943
step: 180, loss: 0.3103128671646118
step: 190, loss: 0.2760572135448456
/usr/local/lib/python3.7/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=0.20162601626016258, f1=0.20501635768811338,
best_f1=0.2056384742951907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24644383788108826
step: 10, loss: 0.09800436347723007
step: 20, loss: 0.29374784231185913
step: 30, loss: 0.16731765866279602
step: 40, loss: 0.30122143030166626
step: 50, loss: 0.33798521757125854
step: 60, loss: 0.3677380681037903
step: 70, loss: 0.12806157767772675
step: 80, loss: 0.3235834836959839
step: 90, loss: 0.8222252726554871
step: 100, loss: 0.3465018570423126
step: 110, loss: 0.3338718116283417
step: 120, loss: 0.578298032283783
step: 130, loss: 0.28159716725349426
step: 140, loss: 0.30895373225212097
step: 150, loss: 0.23366256058216095
step: 160, loss: 0.21232223510742188
step: 170, loss: 0.4940967559814453
step: 180, loss: 0.36249613761901855
step: 190, loss: 0.20113274455070496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24234354194407456, f1=0.19491525423728812,
best_f1=0.19491525423728812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2453760802745819
step: 10, loss: 0.15017685294151306
step: 20, loss: 0.28897014260292053
step: 30, loss: 0.38524681329727173
step: 40, loss: 0.27775973081588745
step: 50, loss: 0.09142208844423294
step: 60, loss: 0.2741999328136444
step: 70, loss: 0.31017962098121643
step: 80, loss: 0.41155514121055603
step: 90, loss: 0.14896339178085327
step: 100, loss: 0.18468479812145233
step: 110, loss: 0.19563178718090057
step: 120, loss: 0.3230625092983246
step: 130, loss: 0.4750067889690399
step: 140, loss: 0.3583833575248718
step: 150, loss: 0.27423375844955444
step: 160, loss: 0.21206310391426086
step: 170, loss: 0.3056667447090149
step: 180, loss: 0.20627905428409576
step: 190, loss: 0.15944594144821167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2657580919931857, f1=0.20598006644518274,
best_f1=0.20598006644518274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4604913890361786
step: 10, loss: 0.09380017220973969
step: 20, loss: 0.2005940079689026
step: 30, loss: 0.18292009830474854
step: 40, loss: 0.07256033271551132
step: 50, loss: 0.312153697013855
step: 60, loss: 0.2067435085773468
step: 70, loss: 0.47050729393959045
step: 80, loss: 0.3610112965106964

```

```

step: 90, loss: 0.42439836263656616
step: 100, loss: 0.17076539993286133
step: 110, loss: 0.24628740549087524
step: 120, loss: 0.33702558279037476
step: 130, loss: 0.5681467056274414
step: 140, loss: 0.46418312191963196
step: 150, loss: 0.2654694616794586
step: 160, loss: 0.2544904351234436
step: 170, loss: 0.39883479475975037
step: 180, loss: 0.319094181060791
step: 190, loss: 0.10977345705032349
/usr/local/lib/python3.7/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=0.25846153846153846, f1=0.25079702444208285,
best_f1=0.20598006644518274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3498089909553528
step: 10, loss: 0.24813570082187653
step: 20, loss: 0.4925702214241028
step: 30, loss: 0.3208407461643219
step: 40, loss: 0.23202453553676605
step: 50, loss: 0.45447826385498047
step: 60, loss: 0.4253377914428711
step: 70, loss: 0.3371959328651428
step: 80, loss: 0.3228009343147278
step: 90, loss: 0.43550023436546326
step: 100, loss: 0.25968289375305176
step: 110, loss: 0.4515277147293091
step: 120, loss: 0.15655383467674255
step: 130, loss: 0.2324681431055069
step: 140, loss: 0.32042551040649414
step: 150, loss: 0.3492547571659088
step: 160, loss: 0.24204151332378387
step: 170, loss: 0.35076603293418884
step: 180, loss: 0.3196199834346771
step: 190, loss: 0.14935480058193207
/usr/local/lib/python3.7/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=0.2656072644721907, f1=0.2541567695961995,
best_f1=0.20598006644518274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27782583236694336

```

```

step: 10, loss: 0.347313791513443
step: 20, loss: 0.32525473833084106
step: 30, loss: 0.3682195544242859
step: 40, loss: 0.21411700546741486
step: 50, loss: 0.21650265157222748
step: 60, loss: 0.2569584846496582
step: 70, loss: 0.1524791419506073
step: 80, loss: 0.2685492932796478
step: 90, loss: 0.272438645362854
step: 100, loss: 0.14256127178668976
step: 110, loss: 0.528336226940155
step: 120, loss: 0.19958478212356567
step: 130, loss: 0.17496339976787567
step: 140, loss: 0.27240216732025146
step: 150, loss: 0.33877161145210266
step: 160, loss: 0.3804394602775574
step: 170, loss: 0.10166409611701965
step: 180, loss: 0.22060847282409668
step: 190, loss: 0.32296356558799744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.27920227920227925, f1=0.24891461649782923,
best_f1=0.24891461649782923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26578831672668457
step: 10, loss: 0.1982097327709198
step: 20, loss: 0.36865630745887756
step: 30, loss: 0.10827303677797318
step: 40, loss: 0.27096009254455566
step: 50, loss: 0.308322548866272
step: 60, loss: 0.2330688238143921
step: 70, loss: 0.20983897149562836
step: 80, loss: 0.3938664197921753
step: 90, loss: 0.18716861307621002
step: 100, loss: 0.2383958250284195
step: 110, loss: 0.4100790023803711
step: 120, loss: 0.2410387247800827
step: 130, loss: 0.3500058054924011
step: 140, loss: 0.17449474334716797
step: 150, loss: 0.32893791794776917
step: 160, loss: 0.29642435908317566

```

```

step: 170, loss: 0.19511587917804718
step: 180, loss: 0.25832709670066833
step: 190, loss: 0.5141288042068481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2807825086306099, f1=0.26265060240963856,
best_f1=0.26265060240963856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5362208485603333
step: 10, loss: 0.19796651601791382
step: 20, loss: 0.2753017544746399
step: 30, loss: 0.2832903563976288
step: 40, loss: 0.5505320429801941
step: 50, loss: 0.29221537709236145
step: 60, loss: 0.20564286410808563
step: 70, loss: 0.19459837675094604
step: 80, loss: 0.18262706696987152
step: 90, loss: 0.29281437397003174
step: 100, loss: 0.2622181177139282
step: 110, loss: 0.36915138363838196
step: 120, loss: 0.19272767007350922
step: 130, loss: 0.26419246196746826
step: 140, loss: 0.4637525975704193
step: 150, loss: 0.13247817754745483
step: 160, loss: 0.2482409030199051
step: 170, loss: 0.7005040645599365
step: 180, loss: 0.2418641448020935
step: 190, loss: 0.3466905355453491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2936416184971098, f1=0.2908224076281287,
best_f1=0.2908224076281287
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 [00:13, 149.93it/s]
load_f1 = 0.28918322295805743
real_f1 = 0.28540305010893247
733it [00:00, 3275.51it/s]/usr/local/lib/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, 136.41it/s]

```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

```

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

```



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

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

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

Step 3: run the matcher

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4819668233394623
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5082944631576538
step: 20, loss: 0.3304861783981323
step: 30, loss: 0.42148470878601074
step: 40, loss: 0.5410178899765015
step: 50, loss: 0.35009777545928955
step: 60, loss: 0.5632655620574951
step: 70, loss: 0.2953180968761444
step: 80, loss: 0.19879402220249176
step: 90, loss: 0.21649880707263947
step: 100, loss: 0.15872855484485626
step: 110, loss: 0.37564817070961
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.31783533096313477
step: 130, loss: 0.31265515089035034
step: 140, loss: 0.3940965533256531
step: 150, loss: 0.3190653324127197
step: 160, loss: 0.3712363541126251
step: 170, loss: 0.3440922498703003
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.3314468562602997

```

```

step: 10, loss: 0.47133761644363403
step: 20, loss: 0.3233257532119751
step: 30, loss: 0.2798701524734497
step: 40, loss: 0.09524879604578018
step: 50, loss: 0.44891276955604553
step: 60, loss: 0.18540668487548828
step: 70, loss: 0.5001113414764404
step: 80, loss: 0.2408686727285385
step: 90, loss: 0.24753208458423615
step: 100, loss: 0.512532651424408
step: 110, loss: 0.25416994094848633
step: 120, loss: 0.2505027949810028
step: 130, loss: 0.5412604808807373
step: 140, loss: 0.5239184498786926
step: 150, loss: 0.43593356013298035
step: 160, loss: 0.44778504967689514
step: 170, loss: 0.37630361318588257
/usr/local/lib/python3.7/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: 0.6081399321556091
step: 10, loss: 0.31332048773765564
step: 20, loss: 0.2318788468837738
step: 30, loss: 0.24745610356330872
step: 40, loss: 0.3763751685619354
step: 50, loss: 0.5780563950538635
step: 60, loss: 0.31425729393959045
step: 70, loss: 0.23585766553878784
step: 80, loss: 0.3740309178829193
step: 90, loss: 0.552177369594574
step: 100, loss: 0.32403674721717834
step: 110, loss: 0.16322743892669678
step: 120, loss: 0.5928425788879395
step: 130, loss: 0.5054849982261658
step: 140, loss: 0.42938876152038574
step: 150, loss: 0.2059735655784607
step: 160, loss: 0.16996492445468903
step: 170, loss: 0.3166041076183319
/usr/local/lib/python3.7/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: 0.3674633204936981
step: 10, loss: 0.6028569936752319
step: 20, loss: 0.2490580976009369
step: 30, loss: 0.45613110065460205
step: 40, loss: 0.24251244962215424
step: 50, loss: 0.38030916452407837
step: 60, loss: 0.7366475462913513
step: 70, loss: 0.3214321732521057
step: 80, loss: 0.45975515246391296
step: 90, loss: 0.3121461272239685
step: 100, loss: 0.3625066578388214
step: 110, loss: 0.447975218296051
step: 120, loss: 0.4912855327129364
step: 130, loss: 0.30797773599624634
step: 140, loss: 0.23604382574558258
step: 150, loss: 0.6961967945098877
step: 160, loss: 0.12500008940696716
step: 170, loss: 0.2464105188846588
/usr/local/lib/python3.7/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: 0.3931814432144165
step: 10, loss: 0.3147989809513092
step: 20, loss: 0.3333466649055481
step: 30, loss: 0.310776948928833
step: 40, loss: 0.253919392824173
step: 50, loss: 0.23728299140930176
step: 60, loss: 0.3133068382740021
step: 70, loss: 0.3101464509963989
step: 80, loss: 0.07884672284126282
step: 90, loss: 0.6183477640151978
step: 100, loss: 0.2681829035282135
step: 110, loss: 0.25486114621162415
step: 120, loss: 0.1255919635295868
step: 130, loss: 0.2410847395658493
step: 140, loss: 0.16936388611793518
step: 150, loss: 0.3069116175174713
step: 160, loss: 0.23809736967086792
step: 170, loss: 0.3669643700122833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/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: 0.3179009258747101
step: 10, loss: 0.10681069642305374
step: 20, loss: 0.381687194108963
step: 30, loss: 0.24339887499809265
step: 40, loss: 0.44543471932411194
step: 50, loss: 0.2569812536239624
step: 60, loss: 0.38779428601264954
step: 70, loss: 0.39678141474723816
step: 80, loss: 0.18842223286628723
step: 90, loss: 0.31540292501449585
step: 100, loss: 0.19178038835525513
step: 110, loss: 0.45050811767578125
step: 120, loss: 0.4456748962402344
step: 130, loss: 0.4828118085861206
step: 140, loss: 0.2544768154621124
step: 150, loss: 0.17581911385059357
step: 160, loss: 0.3107203543186188
step: 170, loss: 0.23623913526535034
/usr/local/lib/python3.7/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=0.1941564561734213
/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: 0.2447480708360672
step: 10, loss: 0.4435291886329651
step: 20, loss: 0.4532339870929718
step: 30, loss: 0.514240562915802
step: 40, loss: 0.12903255224227905
step: 50, loss: 0.43289923667907715
step: 60, loss: 0.558664858341217
step: 70, loss: 0.37872567772865295
step: 80, loss: 0.17515024542808533
step: 90, loss: 0.45585405826568604
step: 100, loss: 0.3081521987915039
step: 110, loss: 0.38208356499671936
step: 120, loss: 0.37299999594688416
step: 130, loss: 0.19414512813091278
step: 140, loss: 0.12601812183856964

```

```

step: 150, loss: 0.31465792655944824
step: 160, loss: 0.3841829001903534
step: 170, loss: 0.3137539327144623
/usr/local/lib/python3.7/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=0.1941564561734213
/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: 0.447349488735199
step: 10, loss: 0.5051223039627075
step: 20, loss: 0.24735093116760254
step: 30, loss: 0.3120020031929016
step: 40, loss: 0.2424827218055725
step: 50, loss: 0.37476977705955505
step: 60, loss: 0.3814496695995331
step: 70, loss: 0.18193098902702332
step: 80, loss: 0.3104762136936188
step: 90, loss: 0.5742186307907104
step: 100, loss: 0.253265380859375
step: 110, loss: 0.5122968554496765
step: 120, loss: 0.32040438055992126
step: 130, loss: 0.31347212195396423
step: 140, loss: 0.5019693970680237
step: 150, loss: 0.2414027899503708
step: 160, loss: 0.18288975954055786
step: 170, loss: 0.45979106426239014
/usr/local/lib/python3.7/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=0.1941564561734213
/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: 0.37930554151535034
step: 10, loss: 0.5826504230499268
step: 20, loss: 0.3790084421634674
step: 30, loss: 0.1805368959903717
step: 40, loss: 0.38380858302116394
step: 50, loss: 0.24983318150043488
step: 60, loss: 0.3730914294719696
step: 70, loss: 0.31065788865089417
step: 80, loss: 0.1648465096950531
step: 90, loss: 0.45446157455444336
step: 100, loss: 0.5010538101196289

```

```

step: 110, loss: 0.31822115182876587
step: 120, loss: 0.44052034616470337
step: 130, loss: 0.31158146262168884
step: 140, loss: 0.37526077032089233
step: 150, loss: 0.3732950985431671
step: 160, loss: 0.24712713062763214
step: 170, loss: 0.43425655364990234
/usr/local/lib/python3.7/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=0.1941564561734213
/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: 0.44435080885887146
step: 10, loss: 0.31531238555908203
step: 20, loss: 0.25172165036201477
step: 30, loss: 0.6407586336135864
step: 40, loss: 0.3712623715400696
step: 50, loss: 0.443731427192688
step: 60, loss: 0.5730729699134827
step: 70, loss: 0.2654813528060913
step: 80, loss: 0.2542800307273865
step: 90, loss: 0.23903034627437592
step: 100, loss: 0.2377023547887802
step: 110, loss: 0.3041459918022156
step: 120, loss: 0.30799150466918945
step: 130, loss: 0.17484726011753082
step: 140, loss: 0.24546360969543457
step: 150, loss: 0.31602928042411804
step: 160, loss: 0.24097774922847748
step: 170, loss: 0.4888654947280884
/usr/local/lib/python3.7/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=0.1941564561734213
/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: 0.3793162405490875
step: 10, loss: 0.3176558017730713
step: 20, loss: 0.1736985445022583
step: 30, loss: 0.5055941343307495
step: 40, loss: 0.24216288328170776
step: 50, loss: 0.24468481540679932
step: 60, loss: 0.3983828127384186

```

```

step: 70, loss: 0.3196139931678772
step: 80, loss: 0.17879533767700195
step: 90, loss: 0.37948334217071533
step: 100, loss: 0.5108690857887268
step: 110, loss: 0.3669043779373169
step: 120, loss: 0.45705172419548035
step: 130, loss: 0.376505047082901
step: 140, loss: 0.5008715987205505
step: 150, loss: 0.4360989034175873
step: 160, loss: 0.25095081329345703
step: 170, loss: 0.18192699551582336
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.226262626262628, f1=0.2318244170096022,
best_f1=0.2318244170096022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38385653495788574
step: 10, loss: 0.29523035883903503
step: 20, loss: 0.23748143017292023
step: 30, loss: 0.2796154022216797
step: 40, loss: 0.3042805790901184
step: 50, loss: 0.3594854772090912
step: 60, loss: 0.2933516800403595
step: 70, loss: 0.16907097399234772
step: 80, loss: 0.24585965275764465
step: 90, loss: 0.3148726224899292
step: 100, loss: 0.25942057371139526
step: 110, loss: 0.4284327030181885
step: 120, loss: 0.3041262626647949
step: 130, loss: 0.37313926219940186
step: 140, loss: 0.2752819359302521
step: 150, loss: 0.28540900349617004
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.7509641647338867
step: 170, loss: 0.3405455946922302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3059210526315789, f1=0.255663430420712,
best_f1=0.255663430420712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3146274983882904
step: 10, loss: 0.19059337675571442
step: 20, loss: 0.2270294576883316
step: 30, loss: 0.44813472032546997
step: 40, loss: 0.37493929266929626
step: 50, loss: 0.1626380980014801
step: 60, loss: 0.22191083431243896
step: 70, loss: 0.31418463587760925
step: 80, loss: 0.18422985076904297
step: 90, loss: 0.2683088779449463
step: 100, loss: 0.13755623996257782
step: 110, loss: 0.41883352398872375
step: 120, loss: 0.10955733805894852
step: 130, loss: 0.12098664790391922
step: 140, loss: 0.18640165030956268
step: 150, loss: 0.3809284269809723
step: 160, loss: 0.058205198496580124
step: 170, loss: 0.08377908915281296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7162790697674418, f1=0.7500000000000001,
best_f1=0.7500000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14539743959903717
step: 10, loss: 0.2357078343629837
step: 20, loss: 0.1644669622182846
step: 30, loss: 0.08253846317529678
step: 40, loss: 0.1283937692642212
step: 50, loss: 0.09513119608163834
step: 60, loss: 0.31934434175491333
step: 70, loss: 0.15614654123783112
step: 80, loss: 0.07233323156833649
step: 90, loss: 0.07217639684677124
step: 100, loss: 0.5204509496688843
step: 110, loss: 0.11073817312717438

```

```

step: 120, loss: 0.19626015424728394
step: 130, loss: 0.20211289823055267
step: 140, loss: 0.06622423976659775
step: 150, loss: 0.11221908032894135
step: 160, loss: 0.18441013991832733
step: 170, loss: 0.16089729964733124
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7632183908045976, f1=0.7845804988662132,
best_f1=0.7845804988662132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024508623406291008
step: 10, loss: 0.13092878460884094
step: 20, loss: 0.22634457051753998
step: 30, loss: 0.2273322343826294
step: 40, loss: 0.03623137250542641
step: 50, loss: 0.054500512778759
step: 60, loss: 0.06729590892791748
step: 70, loss: 0.12521544098854065
step: 80, loss: 0.063688725233078
step: 90, loss: 0.05040152370929718
step: 100, loss: 0.2782520055770874
step: 110, loss: 0.21613483130931854
step: 120, loss: 0.06258600205183029
step: 130, loss: 0.3470067083835602
step: 140, loss: 0.10307107865810394
step: 150, loss: 0.01976454257965088
step: 160, loss: 0.06190973520278931
step: 170, loss: 0.046319734305143356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7704485488126649, f1=0.7846153846153846,
best_f1=0.7846153846153846
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 225.14it/s]
load_f1 = 0.6128133704735377
real_f1 = 0.6293333333333333
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, 137.08it/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/

```

Downloading: 100% 481/481 [00:00<00:00, 423kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 672kB/s]

Downloading: 100% 456k/456k [00:04<00:00, 101kB/s]

Downloading: 100% 501M/501M [00:06<00:00, 72.7MB/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: 0.5610620975494385
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.430954247713089
step: 20, loss: 0.5140743851661682
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.27542760968208313
step: 40, loss: 0.34273990988731384
step: 50, loss: 0.5246987342834473
step: 60, loss: 0.06645911186933517
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.21811118721961975
step: 80, loss: 0.048100706189870834
step: 90, loss: 0.3035450279712677
step: 100, loss: 0.17456784844398499
step: 110, loss: 0.1295962780714035
step: 120, loss: 0.08706557750701904
step: 130, loss: 0.0767948105931282
step: 140, loss: 0.006550635676831007
step: 150, loss: 0.20525068044662476
step: 160, loss: 0.03511791303753853
step: 170, loss: 0.13923417031764984
step: 180, loss: 0.10847140848636627
step: 190, loss: 0.01363984402269125
step: 200, loss: 0.025564901530742645
step: 210, loss: 0.03127725049853325
step: 220, loss: 0.13989976048469543
step: 230, loss: 0.005458390340209007
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9546961325966851, f1=0.9623893805309734,
best_f1=0.9623893805309734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009883681312203407
step: 10, loss: 0.26959946751594543
step: 20, loss: 0.011233880184590816
step: 30, loss: 0.03317427262663841
step: 40, loss: 0.03388640657067299
step: 50, loss: 0.00903208926320076
step: 60, loss: 0.006625481881201267
step: 70, loss: 0.01289411075413227
step: 80, loss: 0.010382524691522121
step: 90, loss: 0.039094820618629456
step: 100, loss: 0.021518534049391747
step: 110, loss: 0.028021566569805145
step: 120, loss: 0.005125239957123995
step: 130, loss: 0.02388060837984085
step: 140, loss: 0.006659325212240219
step: 150, loss: 0.0650186836719513
step: 160, loss: 0.07533200085163116
step: 170, loss: 0.013244034722447395
step: 180, loss: 0.052621182054281235
step: 190, loss: 0.010918114334344864
step: 200, loss: 0.034065138548612595
step: 210, loss: 0.03098728321492672
step: 220, loss: 0.003371114144101739
step: 230, loss: 0.0017007802380248904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9637760702524698, f1=0.9679558011049725,
best_f1=0.9679558011049725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028761662542819977
step: 10, loss: 0.022573431953787804
step: 20, loss: 0.006911423057317734
step: 30, loss: 0.010750651359558105
step: 40, loss: 0.03385516628623009
step: 50, loss: 0.01617206633090973
step: 60, loss: 0.02716580405831337
step: 70, loss: 0.005498893558979034
step: 80, loss: 0.026336541399359703
step: 90, loss: 0.0029762699268758297
step: 100, loss: 0.012680809013545513

```

```

step: 110, loss: 0.008458376862108707
step: 120, loss: 0.021525094285607338
step: 130, loss: 0.0064841099083423615
step: 140, loss: 0.0030771519523113966
step: 150, loss: 0.09086822718381882
step: 160, loss: 0.012653230689466
step: 170, loss: 0.009600980207324028
step: 180, loss: 0.010466977022588253
step: 190, loss: 0.012852521613240242
step: 200, loss: 0.008492397144436836
step: 210, loss: 0.0018939871806651354
step: 220, loss: 0.013575333170592785
step: 230, loss: 0.051920171827077866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753363228699552, f1=0.9617977528089887,
best_f1=0.9617977528089887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009167445823550224
step: 10, loss: 0.00860707089304924
step: 20, loss: 0.006111018359661102
step: 30, loss: 0.007533949799835682
step: 40, loss: 0.06200440600514412
step: 50, loss: 0.010554123669862747
step: 60, loss: 0.003150197444483638
step: 70, loss: 0.020466027781367302
step: 80, loss: 0.10160302370786667
step: 90, loss: 0.06416833400726318
step: 100, loss: 0.012288463301956654
step: 110, loss: 0.002292619552463293
step: 120, loss: 0.025816919282078743
step: 130, loss: 0.011011059395968914
step: 140, loss: 0.02912019193172455
step: 150, loss: 0.003469303948804736
step: 160, loss: 0.0006252505118027329
step: 170, loss: 0.011907643638551235
step: 180, loss: 0.0596812441945076
step: 190, loss: 0.001914511201903224
step: 200, loss: 0.013030214235186577
step: 210, loss: 0.11550282686948776
step: 220, loss: 0.0016103253001347184

```

```

step: 230, loss: 0.005221317522227764
/usr/local/lib/python3.7/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=0.9689578713968958, f1=0.967525195968645,
best_f1=0.9617977528089887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00248138140887022
step: 10, loss: 0.004228959791362286
step: 20, loss: 0.012646044604480267
step: 30, loss: 0.0031105480156838894
step: 40, loss: 0.011697201989591122
step: 50, loss: 0.021570872515439987
step: 60, loss: 0.011643658392131329
step: 70, loss: 0.01727566309273243
step: 80, loss: 0.014355787076056004
step: 90, loss: 0.009413894265890121
step: 100, loss: 0.0007136146887205541
step: 110, loss: 0.0034565322566777468
step: 120, loss: 0.003632000647485256
step: 130, loss: 0.0013433399144560099
step: 140, loss: 0.018822042271494865
step: 150, loss: 0.07542476803064346
step: 160, loss: 0.000380150624550879
step: 170, loss: 0.001750395866110921
step: 180, loss: 0.10941356420516968
step: 190, loss: 0.031077103689312935
step: 200, loss: 0.003914902918040752
step: 210, loss: 0.022839553654193878
step: 220, loss: 0.0014277135487645864
step: 230, loss: 0.07173743844032288
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.976271186440678, f1=0.9751693002257337,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008597150444984436
step: 10, loss: 0.027455730363726616
step: 20, loss: 0.0017150429775938392

```



```

step: 30, loss: 0.0026862621307373047
step: 40, loss: 0.0006637066835537553
step: 50, loss: 0.002925640670582652
step: 60, loss: 0.03803754225373268
step: 70, loss: 0.192081481218338
step: 80, loss: 0.006686212494969368
step: 90, loss: 0.018755868077278137
step: 100, loss: 0.0015840906416997313
step: 110, loss: 0.012051678262650967
step: 120, loss: 0.0008821133524179459
step: 130, loss: 0.0011275879805907607
step: 140, loss: 0.00046008708886802197
step: 150, loss: 0.00053820526227355
step: 160, loss: 0.005966716445982456
step: 170, loss: 0.021253066137433052
step: 180, loss: 0.06793218851089478
step: 190, loss: 0.0004688084591180086
step: 200, loss: 0.013268179260194302
step: 210, loss: 0.000663464074023068
step: 220, loss: 0.04953497648239136
step: 230, loss: 0.001853675814345479
/usr/local/lib/python3.7/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=0.9709821428571428, f1=0.9740698985343857,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005058820825070143
step: 10, loss: 0.0013696375535801053
step: 20, loss: 0.000572866469155997
step: 30, loss: 0.0008055999642238021
step: 40, loss: 0.001710567157715559
step: 50, loss: 0.0019143999088555574
step: 60, loss: 0.0013126203557476401
step: 70, loss: 0.0005321292555890977
step: 80, loss: 0.000470704777399078
step: 90, loss: 0.006954679731279612
step: 100, loss: 0.00034037212026305497
step: 110, loss: 0.00067072146339342
step: 120, loss: 0.00035483177634887397
step: 130, loss: 0.003233653726056218
step: 140, loss: 0.00023893119941931218
step: 150, loss: 0.0015384237049147487
step: 160, loss: 0.00021198684407863766
step: 170, loss: 0.0011032733600586653
step: 180, loss: 0.0002168829960282892

```

```

step: 190, loss: 0.08527717739343643
step: 200, loss: 0.030803393572568893
step: 210, loss: 0.02996818721294403
step: 220, loss: 0.0012909990036860108
step: 230, loss: 0.002407104941084981
/usr/local/lib/python3.7/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=0.9720670391061451, f1=0.9684684684684683,
best_f1=0.9751693002257337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019083119696006179
step: 10, loss: 0.011442933231592178
step: 20, loss: 0.00035093995393253863
step: 30, loss: 0.0006636119796894491
step: 40, loss: 0.0015998382586985826
step: 50, loss: 0.016620084643363953
step: 60, loss: 0.0022901156917214394
step: 70, loss: 0.000736708811018616
step: 80, loss: 0.007640888914465904
step: 90, loss: 0.002356412820518017
step: 100, loss: 0.0009092765394598246
step: 110, loss: 0.0023375260643661022
step: 120, loss: 0.009217026643455029
step: 130, loss: 0.00026962615083903074
step: 140, loss: 0.00018458743579685688
step: 150, loss: 0.07012981176376343
step: 160, loss: 0.012737962417304516
step: 170, loss: 0.11354441195726395
step: 180, loss: 0.0008248916128650308
step: 190, loss: 0.008957911282777786
step: 200, loss: 0.018164262175559998
step: 210, loss: 0.001118397107347846
step: 220, loss: 0.00047508778516203165
step: 230, loss: 0.0008991092327050865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775784753363228, f1=0.9784824462061155,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024224340450018644
step: 10, loss: 0.0007750260992906988
step: 20, loss: 0.0010711231734603643
step: 30, loss: 0.001130552962422371
step: 40, loss: 0.008777972310781479
step: 50, loss: 0.0009526588837616146
step: 60, loss: 0.0007475381135009229
step: 70, loss: 0.04424620792269707
step: 80, loss: 0.0001903298543766141
step: 90, loss: 0.08566946536302567
step: 100, loss: 0.006038610357791185
step: 110, loss: 0.00024799746461212635
step: 120, loss: 0.028649328276515007
step: 130, loss: 0.0021375189535319805
step: 140, loss: 0.0008021958637982607
step: 150, loss: 0.004731275606900454
step: 160, loss: 0.0027860745321959257
step: 170, loss: 0.0013942563673481345
step: 180, loss: 0.00044317994615994394
step: 190, loss: 0.0003594313748180866
step: 200, loss: 0.00036564661422744393
step: 210, loss: 0.01574915461242199
step: 220, loss: 0.0003698163782246411
step: 230, loss: 0.00026272499235346913
/usr/local/lib/python3.7/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=0.9752252252252253, f1=0.9738933030646991,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004297657578717917
step: 10, loss: 0.0033568809740245342
step: 20, loss: 0.000752479478251189
step: 30, loss: 0.0004975594347342849
step: 40, loss: 0.02425774559378624
step: 50, loss: 0.00036142338649369776
step: 60, loss: 0.00027617596788331866
step: 70, loss: 0.0021170219406485558
step: 80, loss: 0.00012107406655559316
step: 90, loss: 0.00012068015348631889
step: 100, loss: 0.00014518265379592776
step: 110, loss: 0.026753762736916542
step: 120, loss: 0.00027204910293221474
step: 130, loss: 0.0002875843201763928
step: 140, loss: 0.00014250972890295088

```

```

step: 150, loss: 0.0005662366165779531
step: 160, loss: 0.00010397216829005629
step: 170, loss: 0.0030775039922446012
step: 180, loss: 0.00133711623493582
step: 190, loss: 0.00029288590303622186
step: 200, loss: 0.0012057459680363536
step: 210, loss: 0.004700841847807169
step: 220, loss: 0.004272400867193937
step: 230, loss: 0.00018113819533027709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9817767653758542, f1=0.9690011481056257,
best_f1=0.9690011481056257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004364659253042191
step: 10, loss: 0.0002896225778385997
step: 20, loss: 0.0002254844584967941
step: 30, loss: 0.0005304139340296388
step: 40, loss: 0.0001547715364722535
step: 50, loss: 0.007203148677945137
step: 60, loss: 0.013541481457650661
step: 70, loss: 0.001161237945780158
step: 80, loss: 0.0034501194022595882
step: 90, loss: 0.012618507258594036
step: 100, loss: 0.00023890180455055088
step: 110, loss: 0.0006282022222876549
step: 120, loss: 0.0010055545717477798
step: 130, loss: 0.00014293658023234457
step: 140, loss: 0.004720584023743868
step: 150, loss: 0.00019334148964844644
step: 160, loss: 0.04035200551152229
step: 170, loss: 0.0012933072866871953
step: 180, loss: 0.0005134472739882767
step: 190, loss: 0.0005691364640370011
step: 200, loss: 0.0037928682286292315
step: 210, loss: 0.0002680057077668607
step: 220, loss: 0.000510798767209053
step: 230, loss: 0.00028956690221093595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9831271091113611, f1=0.9750566893424036,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033870048355311155
step: 10, loss: 0.0004946136032231152
step: 20, loss: 0.010825594887137413
step: 30, loss: 0.06545629352331161
step: 40, loss: 0.0004730010114144534
step: 50, loss: 0.0030867597088217735
step: 60, loss: 0.0007034590817056596
step: 70, loss: 0.00017479373491369188
step: 80, loss: 0.0001265356841031462
step: 90, loss: 0.0014521133853122592
step: 100, loss: 0.00012815797526855022
step: 110, loss: 0.00010350280354032293
step: 120, loss: 0.002696906216442585
step: 130, loss: 0.00011330096458550543
step: 140, loss: 0.00012980549945496023
step: 150, loss: 0.0001646235614316538
step: 160, loss: 0.00014700007159262896
step: 170, loss: 0.00018517229182180017
step: 180, loss: 9.231795411324129e-05
step: 190, loss: 0.0008928753086365759
step: 200, loss: 0.00011964762234129012
step: 210, loss: 0.005921445321291685
step: 220, loss: 0.00036377841024659574
step: 230, loss: 0.0007195041398517787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841628959276018, f1=0.9749430523917996,
best_f1=0.9749430523917996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004243809264153242
step: 10, loss: 0.00016452366253361106
step: 20, loss: 0.00015531173266936094

```

```

step: 30, loss: 0.001015752786770463
step: 40, loss: 0.00013373589899856597
step: 50, loss: 0.00033020315458998084
step: 60, loss: 0.00010803047916851938
step: 70, loss: 0.0008047367446124554
step: 80, loss: 0.0006542385672219098
step: 90, loss: 9.116355067817494e-05
step: 100, loss: 0.00016961197252385318
step: 110, loss: 0.003379751928150654
step: 120, loss: 0.00010454290168127045
step: 130, loss: 0.00011855840421048924
step: 140, loss: 0.00010076074977405369
step: 150, loss: 0.0005737023311667144
step: 160, loss: 0.00013950855645816773
step: 170, loss: 8.273531420854852e-05
step: 180, loss: 0.04683757945895195
step: 190, loss: 8.795916073722765e-05
step: 200, loss: 5.223148764343932e-05
step: 210, loss: 0.00016607956786174327
step: 220, loss: 7.97935645096004e-05
step: 230, loss: 0.00010575849591987208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841986455981941, f1=0.9785310734463276,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011736356827896088
step: 10, loss: 9.103344928007573e-05
step: 20, loss: 8.352949225809425e-05
step: 30, loss: 9.497864812146872e-05
step: 40, loss: 9.239062637789175e-05
step: 50, loss: 4.48972059530206e-05
step: 60, loss: 6.753925117664039e-05
step: 70, loss: 0.00028997453046031296
step: 80, loss: 5.137399784871377e-05
step: 90, loss: 0.00014331469719763845
step: 100, loss: 8.05492527433671e-05
step: 110, loss: 0.0014053593622520566
step: 120, loss: 3.3890788472490385e-05
step: 130, loss: 0.0001535168703412637
step: 140, loss: 0.00021973319235257804

```

```

step: 150, loss: 3.215459219063632e-05
step: 160, loss: 0.00011851742601720616
step: 170, loss: 0.00024528210633434355
step: 180, loss: 0.00024913009838201106
step: 190, loss: 0.00013868419046048075
step: 200, loss: 0.0004474601009860635
step: 210, loss: 8.985819295048714e-05
step: 220, loss: 7.854992873035371e-05
step: 230, loss: 0.0002837258798535913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842342342342343, f1=0.9796839729119639,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04046402499079704
step: 10, loss: 0.00039500423008576035
step: 20, loss: 0.00011858048674184829
step: 30, loss: 8.439599332632497e-05
step: 40, loss: 5.3956013289280236e-05
step: 50, loss: 0.0001027741891448386
step: 60, loss: 0.02552446722984314
step: 70, loss: 0.0028094169683754444
step: 80, loss: 7.434383587678894e-05
step: 90, loss: 5.535249874810688e-05
step: 100, loss: 5.914177017984912e-05
step: 110, loss: 0.000974498107098043
step: 120, loss: 0.02513054944574833
step: 130, loss: 0.0031423273030668497
step: 140, loss: 0.0003189142735209316
step: 150, loss: 0.00021254913008306175
step: 160, loss: 0.01932465471327305
step: 170, loss: 0.000480469228932634
step: 180, loss: 8.337973849847913e-05
step: 190, loss: 0.00022076955065131187
step: 200, loss: 0.0004012783756479621
step: 210, loss: 0.03154318407177925
step: 220, loss: 0.00011053606431232765
step: 230, loss: 0.0001550772722112015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853768278965129, f1=0.9796839729119639,
best_f1=0.9796839729119639
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

```
2473it [00:16, 154.06it/s]
```

```
load_f1 = 0.983050847457627
```

```
real_f1 = 0.9830890642615557
```

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

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

```
4398it [00:30, 142.48it/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: 0.6149146556854248
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: 0.37048614025115967
step: 20, loss: 0.32321396470069885
step: 30, loss: 0.41160696744918823
step: 40, loss: 0.36790943145751953
step: 50, loss: 0.5762134790420532
step: 60, loss: 0.4012279212474823
step: 70, loss: 0.5705980658531189
step: 80, loss: 0.454135924577713
step: 90, loss: 0.21020159125328064
step: 100, loss: 0.3319016695022583
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: 110, loss: 0.13558802008628845
step: 120, loss: 0.3617953658103943
step: 130, loss: 0.15167774260044098
step: 140, loss: 0.3159269392490387
step: 150, loss: 0.17115367949008942
step: 160, loss: 0.26486438512802124
step: 170, loss: 0.07483869791030884
step: 180, loss: 0.06923886388540268
step: 190, loss: 0.0974944531917572
step: 200, loss: 0.12998680770397186
step: 210, loss: 0.0395134799182415
step: 220, loss: 0.1047617644071579
step: 230, loss: 0.15806081891059875
step: 240, loss: 0.05936376377940178
step: 250, loss: 0.03925534710288048
step: 260, loss: 0.3248765170574188
step: 270, loss: 0.16548578441143036
step: 280, loss: 0.07729741185903549

```

```

step: 290, loss: 0.15058167278766632
step: 300, loss: 0.03681175783276558
step: 310, loss: 0.12422551214694977
step: 320, loss: 0.0649518221616745
step: 330, loss: 0.047720205038785934
step: 340, loss: 0.4042102098464966
step: 350, loss: 0.10581620782613754
step: 360, loss: 0.027014387771487236
step: 370, loss: 0.03210932016372681
step: 380, loss: 0.1830001324415207
step: 390, loss: 0.025337237864732742
step: 400, loss: 0.02359476312994957
step: 410, loss: 0.24995043873786926
step: 420, loss: 0.029510021209716797
step: 430, loss: 0.027528945356607437
step: 440, loss: 0.07609762996435165
step: 450, loss: 0.03810938447713852
step: 460, loss: 0.07773394137620926
step: 470, loss: 0.04110444337129593
step: 480, loss: 0.10374816507101059
step: 490, loss: 0.32428961992263794
step: 500, loss: 0.017511039972305298
step: 510, loss: 0.07266008108854294
step: 520, loss: 0.0851893201470375
step: 530, loss: 0.013893461786210537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.933579335793358, f1=0.9348729792147806,
best_f1=0.9348729792147806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09996750205755234
step: 10, loss: 0.06461764127016068
step: 20, loss: 0.021678607910871506
step: 30, loss: 0.18688975274562836
step: 40, loss: 0.10867141932249069
step: 50, loss: 0.04241019859910011
step: 60, loss: 0.03359553962945938
step: 70, loss: 0.03615613281726837
step: 80, loss: 0.007973221130669117
step: 90, loss: 0.024246517568826675
step: 100, loss: 0.08023350685834885

```

```
step: 110, loss: 0.006417888216674328
step: 120, loss: 0.04259994626045227
step: 130, loss: 0.006736373528838158
step: 140, loss: 0.04530206322669983
step: 150, loss: 0.026899639517068863
step: 160, loss: 0.025298859924077988
step: 170, loss: 0.026738498359918594
step: 180, loss: 0.02534114569425583
step: 190, loss: 0.07827972620725632
step: 200, loss: 0.26674774289131165
step: 210, loss: 0.04109738767147064
step: 220, loss: 0.0054700965993106365
step: 230, loss: 0.10526023060083389
step: 240, loss: 0.005850617308169603
step: 250, loss: 0.11470288038253784
step: 260, loss: 0.01770126260817051
step: 270, loss: 0.016838116571307182
step: 280, loss: 0.11214133352041245
step: 290, loss: 0.06823026388883591
step: 300, loss: 0.04357339069247246
step: 310, loss: 0.028391582891345024
step: 320, loss: 0.037370722740888596
step: 330, loss: 0.037657495588064194
step: 340, loss: 0.16256821155548096
step: 350, loss: 0.002801097696647048
step: 360, loss: 0.07705260813236237
step: 370, loss: 0.026558425277471542
step: 380, loss: 0.1757914125919342
step: 390, loss: 0.015318808145821095
step: 400, loss: 0.040738753974437714
step: 410, loss: 0.055548712611198425
step: 420, loss: 0.16375049948692322
step: 430, loss: 0.13564617931842804
step: 440, loss: 0.030113479122519493
step: 450, loss: 0.11272721737623215
step: 460, loss: 0.0333683155477047
step: 470, loss: 0.02816881239414215
step: 480, loss: 0.01914713904261589
step: 490, loss: 0.07025297731161118
step: 500, loss: 0.008007139898836613
step: 510, loss: 0.046219754964113235
step: 520, loss: 0.23035667836666107
step: 530, loss: 0.11469404399394989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9458082445576655, f1=0.9385113268608415,
best_f1=0.9385113268608415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11040682345628738
step: 10, loss: 0.036583442240953445
step: 20, loss: 0.017104320228099823
step: 30, loss: 0.03004266507923603
step: 40, loss: 0.28167402744293213
step: 50, loss: 0.004746992141008377
step: 60, loss: 0.010507967323064804
step: 70, loss: 0.022001156583428383
step: 80, loss: 0.21277636289596558
step: 90, loss: 0.0810089111328125
step: 100, loss: 0.029980046674609184
step: 110, loss: 0.025595763698220253
step: 120, loss: 0.22270125150680542
step: 130, loss: 0.07439468055963516
step: 140, loss: 0.014060878194868565
step: 150, loss: 0.04442329332232475
step: 160, loss: 0.01858900859951973
step: 170, loss: 0.007059051655232906
step: 180, loss: 0.014272480271756649
step: 190, loss: 0.00524531863629818
step: 200, loss: 0.03615095093846321
step: 210, loss: 0.06470004469156265
step: 220, loss: 0.10624220222234726
step: 230, loss: 0.04373599588871002
step: 240, loss: 0.061998751014471054
step: 250, loss: 0.08952081203460693
step: 260, loss: 0.04343194141983986
step: 270, loss: 0.0097477026283741
step: 280, loss: 0.08627980947494507
step: 290, loss: 0.032706718891859055
step: 300, loss: 0.03737391531467438
step: 310, loss: 0.11701931804418564
step: 320, loss: 0.012512445449829102
step: 330, loss: 0.005968870595097542
step: 340, loss: 0.0100303515791893
step: 350, loss: 0.04893994703888893
step: 360, loss: 0.006573984865099192
step: 370, loss: 0.09402123838663101
step: 380, loss: 0.0031302371062338352
step: 390, loss: 0.023933347314596176
step: 400, loss: 0.18690523505210876

```

```

step: 410, loss: 0.008351118303835392
step: 420, loss: 0.012392650358378887
step: 430, loss: 0.027126604691147804
step: 440, loss: 0.3917285203933716
step: 450, loss: 0.026070155203342438
step: 460, loss: 0.0749126672744751
step: 470, loss: 0.0384896956384182
step: 480, loss: 0.07099904119968414
step: 490, loss: 0.005853260867297649
step: 500, loss: 0.08032682538032532
step: 510, loss: 0.018879912793636322
step: 520, loss: 0.028399556875228882
step: 530, loss: 0.00853561982512474
/usr/local/lib/python3.7/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=0.9422263973696571, f1=0.9347826086956521,
best_f1=0.9385113268608415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03221851587295532
step: 10, loss: 0.015603815205395222
step: 20, loss: 0.09595340490341187
step: 30, loss: 0.028314698487520218
step: 40, loss: 0.002648942871019244
step: 50, loss: 0.07948268204927444
step: 60, loss: 0.011493842117488384
step: 70, loss: 0.20738717913627625
step: 80, loss: 0.1309671550989151
step: 90, loss: 0.047996945679187775
step: 100, loss: 0.025402221828699112
step: 110, loss: 0.20104491710662842
step: 120, loss: 0.005695607513189316
step: 130, loss: 0.015198170207440853
step: 140, loss: 0.0605287104845047
step: 150, loss: 0.08302759379148483
step: 160, loss: 0.004825320560485125
step: 170, loss: 0.018915699794888496
step: 180, loss: 0.016765624284744263
step: 190, loss: 0.04679407924413681
step: 200, loss: 0.03361188620328903
step: 210, loss: 0.0005443575792014599
step: 220, loss: 0.0014254390262067318
step: 230, loss: 0.01844940148293972
step: 240, loss: 0.027611779049038887
step: 250, loss: 0.08518551290035248
step: 260, loss: 0.0036315557081252337

```

```

step: 270, loss: 0.029032528400421143
step: 280, loss: 0.002166267717257142
step: 290, loss: 0.011370968073606491
step: 300, loss: 0.0022070773411542177
step: 310, loss: 0.0036215283907949924
step: 320, loss: 0.021229926496744156
step: 330, loss: 0.025650186464190483
step: 340, loss: 0.004700613208115101
step: 350, loss: 0.042695339769124985
step: 360, loss: 0.09858176857233047
step: 370, loss: 0.008663673885166645
step: 380, loss: 0.006820142734795809
step: 390, loss: 0.002255114493891597
step: 400, loss: 0.007779676001518965
step: 410, loss: 0.020024076104164124
step: 420, loss: 0.012269662693142891
step: 430, loss: 0.02971726842224598
step: 440, loss: 0.032280001789331436
step: 450, loss: 0.019435647875070572
step: 460, loss: 0.0034803326707333326
step: 470, loss: 0.0019238274544477463
step: 480, loss: 0.03444177657365799
step: 490, loss: 0.04418139532208443
step: 500, loss: 0.04994890093803406
step: 510, loss: 0.07294896990060806
step: 520, loss: 0.011061903089284897
step: 530, loss: 0.12113720178604126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9471715755025713, f1=0.9387186629526463,
best_f1=0.9387186629526463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021317297592759132
step: 10, loss: 0.013672389090061188
step: 20, loss: 0.019078494980931282
step: 30, loss: 0.0020507576409727335
step: 40, loss: 0.0008604691247455776
step: 50, loss: 0.00870898924767971
step: 60, loss: 0.0053592948243021965
step: 70, loss: 0.0016391624230891466
step: 80, loss: 0.010825833305716515

```

```
step: 90, loss: 0.02356080338358879
step: 100, loss: 0.09150346368551254
step: 110, loss: 0.04358268529176712
step: 120, loss: 0.06851492077112198
step: 130, loss: 0.047973498702049255
step: 140, loss: 0.00455094687640667
step: 150, loss: 0.0428510457277298
step: 160, loss: 0.03860443830490112
step: 170, loss: 0.04068596661090851
step: 180, loss: 0.04618861898779869
step: 190, loss: 0.006334200501441956
step: 200, loss: 0.009361298754811287
step: 210, loss: 0.0011321342317387462
step: 220, loss: 0.010491423308849335
step: 230, loss: 0.002499983413144946
step: 240, loss: 0.041484713554382324
step: 250, loss: 0.12533700466156006
step: 260, loss: 0.002547109965234995
step: 270, loss: 0.0025976577308028936
step: 280, loss: 0.050970032811164856
step: 290, loss: 0.0021434975787997246
step: 300, loss: 0.08341429382562637
step: 310, loss: 0.02646363526582718
step: 320, loss: 0.01801822893321514
step: 330, loss: 0.023430464789271355
step: 340, loss: 0.033698756247758865
step: 350, loss: 0.04587450623512268
step: 360, loss: 0.00040978979086503386
step: 370, loss: 0.0007854431751184165
step: 380, loss: 0.0004597888619173318
step: 390, loss: 0.0028764635790139437
step: 400, loss: 0.07795874029397964
step: 410, loss: 0.11894306540489197
step: 420, loss: 0.11486788839101791
step: 430, loss: 0.0025628595612943172
step: 440, loss: 0.002677831333130598
step: 450, loss: 0.005534104071557522
step: 460, loss: 0.06525570899248123
step: 470, loss: 0.047367360442876816
step: 480, loss: 0.03288928046822548
step: 490, loss: 0.09494378417730331
step: 500, loss: 0.024636082351207733
step: 510, loss: 0.007604556158185005
step: 520, loss: 0.08811990171670914
step: 530, loss: 0.014752744697034359
/usr/local/lib/python3.7/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=0.9462068965517241, f1=0.933271547729379,  
best_f1=0.9387186629526463  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.11896153539419174  
step: 10, loss: 0.00947427935898304  
step: 20, loss: 0.009938899427652359  
step: 30, loss: 0.0011242673499509692  
step: 40, loss: 0.0018598514143377542  
step: 50, loss: 0.004948865156620741  
step: 60, loss: 0.012945941649377346  
step: 70, loss: 0.0027748337015509605  
step: 80, loss: 0.07595986127853394  
step: 90, loss: 0.0020993100479245186  
step: 100, loss: 0.034506551921367645  
step: 110, loss: 0.008281100541353226  
step: 120, loss: 0.08585502952337265  
step: 130, loss: 0.018226850777864456  
step: 140, loss: 0.0038777452427893877  
step: 150, loss: 0.0013057123869657516  
step: 160, loss: 0.07705943286418915  
step: 170, loss: 0.00625836756080389  
step: 180, loss: 0.01850113645195961  
step: 190, loss: 0.314392626285553  
step: 200, loss: 0.02250088006258011  
step: 210, loss: 0.015305657871067524  
step: 220, loss: 0.0037792471703141928  
step: 230, loss: 0.014236421324312687  
step: 240, loss: 0.01829446293413639  
step: 250, loss: 0.06194530054926872  
step: 260, loss: 0.0011366767575964332  
step: 270, loss: 0.00771399075165391  
step: 280, loss: 0.0039055643137544394  
step: 290, loss: 0.005733022931963205  
step: 300, loss: 0.13112467527389526  
step: 310, loss: 0.05461474880576134  
step: 320, loss: 0.0005030310712754726  
step: 330, loss: 0.05615857616066933  
step: 340, loss: 0.0006900735315866768  
step: 350, loss: 0.0006109054083935916  
step: 360, loss: 0.20822666585445404  
step: 370, loss: 0.0049751317128539085  
step: 380, loss: 0.005973485764116049  
step: 390, loss: 0.000886106863617897  
step: 400, loss: 0.004203359130769968  
step: 410, loss: 7.298494165297598e-05  
step: 420, loss: 0.0035460288636386395
```

```
step: 430, loss: 0.0004250782949384302
step: 440, loss: 0.0016250911867246032
step: 450, loss: 0.09673330187797546
step: 460, loss: 0.0012061805464327335
step: 470, loss: 0.0014653595862910151
step: 480, loss: 0.0031273080967366695
step: 490, loss: 0.032933201640844345
step: 500, loss: 0.0037084086798131466
step: 510, loss: 0.12587055563926697
step: 520, loss: 0.0005671812104992568
step: 530, loss: 0.029968073591589928
/usr/local/lib/python3.7/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=0.9405255878284925, f1=0.9297197978870005,
best_f1=0.9387186629526463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04969286546111107
step: 10, loss: 0.0014475365169346333
step: 20, loss: 0.04314536228775978
step: 30, loss: 0.01202332228422165
step: 40, loss: 0.007086253259330988
step: 50, loss: 0.007627982646226883
step: 60, loss: 0.0018362922128289938
step: 70, loss: 0.014482775703072548
step: 80, loss: 0.0019323044689372182
step: 90, loss: 0.008305399678647518
step: 100, loss: 0.0015955225098878145
step: 110, loss: 0.000749966420698911
step: 120, loss: 0.0037090752739459276
step: 130, loss: 0.00011629953951342031
step: 140, loss: 0.0022220781538635492
step: 150, loss: 0.0013095370959490538
step: 160, loss: 0.00018421180720906705
step: 170, loss: 0.005179708823561668
step: 180, loss: 0.10952174663543701
step: 190, loss: 0.03948710113763809
step: 200, loss: 0.001589589985087514
step: 210, loss: 0.06165158003568649
step: 220, loss: 0.0009356860537081957
step: 230, loss: 0.071540467441082
step: 240, loss: 0.017157528549432755
step: 250, loss: 0.004004678688943386
step: 260, loss: 0.042231447994709015
step: 270, loss: 0.05979641526937485
step: 280, loss: 0.00978123676031828
```

```

step: 290, loss: 0.0008144705789163709
step: 300, loss: 0.00025338990963064134
step: 310, loss: 0.0014514746144413948
step: 320, loss: 0.0184024665504694
step: 330, loss: 0.0023921802639961243
step: 340, loss: 0.024266688153147697
step: 350, loss: 0.0037876025307923555
step: 360, loss: 0.010986345820128918
step: 370, loss: 0.10628094524145126
step: 380, loss: 0.4395420253276825
step: 390, loss: 0.017672734335064888
step: 400, loss: 0.07414772361516953
step: 410, loss: 0.009710468351840973
step: 420, loss: 0.023204850032925606
step: 430, loss: 0.014657208696007729
step: 440, loss: 0.000631182745564729
step: 450, loss: 0.0007865637307986617
step: 460, loss: 0.0076106879860162735
step: 470, loss: 0.14120464026927948
step: 480, loss: 0.0023430834989994764
step: 490, loss: 0.03811946138739586
step: 500, loss: 0.0004244732845108956
step: 510, loss: 0.0018391278572380543
step: 520, loss: 0.0006144606159068644
step: 530, loss: 0.002639038721099496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9506057781919852, f1=0.9340813464235624,
best_f1=0.9340813464235624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002132877561962232
step: 10, loss: 0.0009755277424119413
step: 20, loss: 0.002163355238735676
step: 30, loss: 0.004684266168624163
step: 40, loss: 4.7670560888946056e-05
step: 50, loss: 0.00034147006226703525
step: 60, loss: 0.0013168886071071029
step: 70, loss: 0.0004509449645411223
step: 80, loss: 0.0016122572124004364
step: 90, loss: 0.0019663905259221792
step: 100, loss: 0.0028919668402522802

```

```
step: 110, loss: 0.0005764457164332271
step: 120, loss: 0.001484600710682571
step: 130, loss: 0.0007208209717646241
step: 140, loss: 0.0003678470675367862
step: 150, loss: 0.010386061854660511
step: 160, loss: 0.0005670481477864087
step: 170, loss: 0.13340993225574493
step: 180, loss: 0.00023885091650299728
step: 190, loss: 0.0005311642889864743
step: 200, loss: 0.0015890449285507202
step: 210, loss: 0.20128972828388214
step: 220, loss: 0.005530728958547115
step: 230, loss: 0.034681420773267746
step: 240, loss: 0.05187563598155975
step: 250, loss: 0.005673292558640242
step: 260, loss: 0.004524268209934235
step: 270, loss: 0.0022943750955164433
step: 280, loss: 0.0005796749610453844
step: 290, loss: 0.026383861899375916
step: 300, loss: 0.00018618206377141178
step: 310, loss: 0.0014433974865823984
step: 320, loss: 6.800903793191537e-05
step: 330, loss: 0.00046692066825926304
step: 340, loss: 0.005032409448176622
step: 350, loss: 0.023281650617718697
step: 360, loss: 0.01597648486495018
step: 370, loss: 0.057688962668180466
step: 380, loss: 0.00033958599669858813
step: 390, loss: 0.004515867214649916
step: 400, loss: 0.036615900695323944
step: 410, loss: 0.027852389961481094
step: 420, loss: 0.00028429957455955446
step: 430, loss: 0.0008091387571766973
step: 440, loss: 0.0003003249003086239
step: 450, loss: 0.00014718469174113125
step: 460, loss: 0.002585948444902897
step: 470, loss: 0.05825360491871834
step: 480, loss: 0.02887088805437088
step: 490, loss: 0.0017492541810497642
step: 500, loss: 0.0038616631645709276
step: 510, loss: 0.017022114247083664
step: 520, loss: 0.0016969849821180105
step: 530, loss: 0.030093906447291374
/usr/local/lib/python3.7/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=0.9471715755025713, f1=0.9416126042632066,
best_f1=0.9340813464235624
```

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

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

step: 0, loss: 0.007431962061673403
step: 10, loss: 0.006061313208192587
step: 20, loss: 0.00021020795975346118
step: 30, loss: 0.04775696247816086
step: 40, loss: 5.037080336478539e-05
step: 50, loss: 9.729204612085596e-05
step: 60, loss: 0.0001154470955953002
step: 70, loss: 0.015308587811887264
step: 80, loss: 0.0005100471316836774
step: 90, loss: 0.11156589537858963
step: 100, loss: 0.0002471848565619439
step: 110, loss: 0.013644066639244556
step: 120, loss: 0.01748546212911606
step: 130, loss: 0.0010945475660264492
step: 140, loss: 0.0007004266954027116
step: 150, loss: 0.001771922456100583
step: 160, loss: 0.011550264433026314
step: 170, loss: 0.006125576328486204
step: 180, loss: 0.04854150861501694
step: 190, loss: 0.026689261198043823
step: 200, loss: 0.00029114363132975996
step: 210, loss: 0.007602598052471876
step: 220, loss: 0.0010491784196346998
step: 230, loss: 9.012540976982564e-05
step: 240, loss: 0.0025189078878611326
step: 250, loss: 0.01014406606554985
step: 260, loss: 0.00012184240767965093
step: 270, loss: 0.012542981654405594
step: 280, loss: 0.00649716192856431
step: 290, loss: 0.0015753997722640634
step: 300, loss: 0.0005192464450374246
step: 310, loss: 0.026507368311285973
step: 320, loss: 0.00031774831586517394
step: 330, loss: 0.0010258262045681477
step: 340, loss: 0.0045887259766459465
step: 350, loss: 0.10749781876802444
step: 360, loss: 0.023883331567049026
step: 370, loss: 0.0018842399585992098
step: 380, loss: 0.043480709195137024
step: 390, loss: 0.00590889248996973
step: 400, loss: 0.030372444540262222
step: 410, loss: 0.0009592183632776141
step: 420, loss: 0.0005239195306785405
step: 430, loss: 0.026765871793031693
step: 440, loss: 4.191350672044791e-05

```

step: 450, loss: 0.0004383880877867341
step: 460, loss: 9.912654059007764e-05
step: 470, loss: 7.459695189027116e-05
step: 480, loss: 0.000347844761563465
step: 490, loss: 0.0009544034255668521
step: 500, loss: 0.00265506561845541
step: 510, loss: 0.0014572748914361
step: 520, loss: 0.005029766354709864
step: 530, loss: 0.03718110918998718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.950668510834486, f1=0.9380449747590638,
best_f1=0.9380449747590638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000634077878203243
step: 10, loss: 0.0007678703404963017
step: 20, loss: 0.0034366301260888577
step: 30, loss: 0.0006709540612064302
step: 40, loss: 0.00013371629756875336
step: 50, loss: 0.0001442816574126482
step: 60, loss: 0.00076869415352121
step: 70, loss: 0.0003219509089831263
step: 80, loss: 0.0030750734731554985
step: 90, loss: 0.0014365717070177197
step: 100, loss: 0.000801276764832437
step: 110, loss: 0.005925323348492384
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.19300587475299835
step: 130, loss: 0.001928783836774528
step: 140, loss: 0.001240733079612255
step: 150, loss: 0.0021802030969411135
step: 160, loss: 0.03711390495300293
step: 170, loss: 0.0003156749880872667
step: 180, loss: 0.010088055394589901
step: 190, loss: 0.00019770915969274938
step: 200, loss: 0.00036592819378711283
step: 210, loss: 0.0007621193071827292
step: 220, loss: 0.0009250015136785805
step: 230, loss: 0.0012746861902996898
step: 240, loss: 0.00023518914531450719
step: 250, loss: 0.0014338516630232334

```

```

step: 260, loss: 0.0012665257090702653
step: 270, loss: 0.00034381638397462666
step: 280, loss: 0.0014943305868655443
step: 290, loss: 0.0005398361827246845
step: 300, loss: 0.002645441796630621
step: 310, loss: 0.03742494061589241
step: 320, loss: 0.0018871568609029055
step: 330, loss: 0.0007573999464511871
step: 340, loss: 0.0004039780469611287
step: 350, loss: 0.000786574266385287
step: 360, loss: 2.9034143153694458e-05
step: 370, loss: 0.002856515347957611
step: 380, loss: 0.0022138357162475586
step: 390, loss: 0.0005546686588786542
step: 400, loss: 0.00027392100309953094
step: 410, loss: 0.0028409603983163834
step: 420, loss: 0.0004655096272472292
step: 430, loss: 0.0005322038778103888
step: 440, loss: 9.268067515222356e-05
step: 450, loss: 0.0008085679146461189
step: 460, loss: 0.0019620233215391636
step: 470, loss: 0.03079114481806755
step: 480, loss: 0.001441243221051991
step: 490, loss: 0.0415210947394371
step: 500, loss: 0.002275102771818638
step: 510, loss: 0.00026693817926570773
step: 520, loss: 0.04473176598548889
step: 530, loss: 0.01282251626253128
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9530450953045096, f1=0.9444444444444445,
best_f1=0.9444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016518362099304795
step: 10, loss: 0.00019152104505337775
step: 20, loss: 0.00035366031806916
step: 30, loss: 0.0009505244088359177
step: 40, loss: 0.00042955452227033675
step: 50, loss: 0.000232727761651203
step: 60, loss: 0.003900860669091344
step: 70, loss: 0.00014490995090454817

```

```
step: 80, loss: 0.003489074297249317
step: 90, loss: 0.002306508831679821
step: 100, loss: 0.0008105012821033597
step: 110, loss: 0.001698292326182127
step: 120, loss: 0.005200750194489956
step: 130, loss: 0.0007551565067842603
step: 140, loss: 0.0010997502831742167
step: 150, loss: 0.004254885949194431
step: 160, loss: 0.0029742619954049587
step: 170, loss: 0.0004929943243041635
step: 180, loss: 0.00016519465134479105
step: 190, loss: 0.001645370153710246
step: 200, loss: 0.0002863496192730963
step: 210, loss: 0.00032479967921972275
step: 220, loss: 0.015772158280014992
step: 230, loss: 6.316410872386768e-05
step: 240, loss: 0.0011894999770447612
step: 250, loss: 9.034208778757602e-05
step: 260, loss: 0.0018747596768662333
step: 270, loss: 0.0029922404792159796
step: 280, loss: 0.0006958087324164808
step: 290, loss: 0.00015087724023032933
step: 300, loss: 0.0012666789116337895
step: 310, loss: 0.0033642700873315334
step: 320, loss: 0.0021870487835258245
step: 330, loss: 3.0336659619933926e-05
step: 340, loss: 0.0031102292705327272
step: 350, loss: 0.0017288713715970516
step: 360, loss: 0.0001625916047487408
step: 370, loss: 0.0004020347259938717
step: 380, loss: 0.00020599467097781599
step: 390, loss: 0.0016329006757587194
step: 400, loss: 0.014262910932302475
step: 410, loss: 0.005766941234469414
step: 420, loss: 0.0021485157776623964
step: 430, loss: 0.0032361806370317936
step: 440, loss: 0.001606883597560227
step: 450, loss: 0.0020681985188275576
step: 460, loss: 0.002509994665160775
step: 470, loss: 0.0019948307890444994
step: 480, loss: 0.00038612785283476114
step: 490, loss: 0.0010529012652114034
step: 500, loss: 0.0020684467162936926
step: 510, loss: 0.12190890312194824
step: 520, loss: 0.0015155471628531814
step: 530, loss: 0.012204539962112904
/usr/local/lib/python3.7/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=0.951501154734411, f1=0.9410138248847926,
best_f1=0.9444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016229255124926567
step: 10, loss: 0.006343742832541466
step: 20, loss: 0.00834632571786642
step: 30, loss: 0.030465658754110336
step: 40, loss: 0.0029280143789947033
step: 50, loss: 0.04381750524044037
step: 60, loss: 0.0018338720547035336
step: 70, loss: 0.011594177223742008
step: 80, loss: 0.0005921798292547464
step: 90, loss: 0.005054776091128588
step: 100, loss: 0.004837649408727884
step: 110, loss: 0.0012757846852764487
step: 120, loss: 0.0004572017933242023
step: 130, loss: 0.006465218961238861
step: 140, loss: 0.0008961822604760528
step: 150, loss: 5.730338307330385e-05
step: 160, loss: 0.0002728812978602946
step: 170, loss: 0.00011497143714223057
step: 180, loss: 0.0001298597053391859
step: 190, loss: 0.000351826281985268
step: 200, loss: 0.001874408801086247
step: 210, loss: 7.499246567022055e-05
step: 220, loss: 0.00011571315553737804
step: 230, loss: 0.0007019918411970139
step: 240, loss: 0.0031666401773691177
step: 250, loss: 5.6634580687386915e-05
step: 260, loss: 0.00036496814573183656
step: 270, loss: 0.023342369124293327
step: 280, loss: 0.00016928957484196872
step: 290, loss: 0.00028202502289786935
step: 300, loss: 0.0005667543155141175
step: 310, loss: 0.003216532524675131
step: 320, loss: 0.0001997118815779686
step: 330, loss: 0.020161785185337067
step: 340, loss: 0.002490856684744358
step: 350, loss: 0.00028792256489396095
step: 360, loss: 5.3506038966588676e-05
step: 370, loss: 0.0019524061353877187
step: 380, loss: 0.0004478069895412773
step: 390, loss: 0.0026799533516168594
step: 400, loss: 0.00022886636725161225
step: 410, loss: 0.0013927598483860493
```

```

step: 420, loss: 0.004086476285010576
step: 430, loss: 0.0007713983068242669
step: 440, loss: 0.000344862142810598
step: 450, loss: 0.006634116172790527
step: 460, loss: 2.3729173335595988e-05
step: 470, loss: 0.002262897090986371
step: 480, loss: 0.05414436757564545
step: 490, loss: 4.322417589719407e-05
step: 500, loss: 5.20563407917507e-05
step: 510, loss: 0.0013697422109544277
step: 520, loss: 0.00036458944668993354
step: 530, loss: 0.0012088080402463675
/usr/local/lib/python3.7/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=0.9526462395543175, f1=0.9365446966188051,
best_f1=0.9444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014264110941439867
step: 10, loss: 0.006087945774197578
step: 20, loss: 0.003391778562217951
step: 30, loss: 0.000408013176638633
step: 40, loss: 0.000500185415148735
step: 50, loss: 0.0012857945403084159
step: 60, loss: 0.00027166021754965186
step: 70, loss: 0.01954314112663269
step: 80, loss: 0.01879965327680111
step: 90, loss: 0.00011153166997246444
step: 100, loss: 0.0009683422977104783
step: 110, loss: 0.0012800248805433512
step: 120, loss: 0.005289692431688309
step: 130, loss: 0.0006655930774286389
step: 140, loss: 0.00021607983217108995
step: 150, loss: 0.001241170335561037
step: 160, loss: 0.0025299203116446733
step: 170, loss: 0.003245402593165636
step: 180, loss: 0.0002187907521147281
step: 190, loss: 0.0001551575114717707
step: 200, loss: 0.001931200735270977
step: 210, loss: 1.8190115952165797e-05
step: 220, loss: 3.216946424799971e-05
step: 230, loss: 0.0006406032480299473
step: 240, loss: 0.000497330620419234
step: 250, loss: 0.0003685545816551894
step: 260, loss: 0.07824977487325668
step: 270, loss: 0.031091617420315742

```

```

step: 280, loss: 0.00032860456849448383
step: 290, loss: 0.00015065597835928202
step: 300, loss: 1.4837379239907023e-05
step: 310, loss: 0.00020110775949433446
step: 320, loss: 5.974302621325478e-05
step: 330, loss: 0.0011482246918603778
step: 340, loss: 0.004447094164788723
step: 350, loss: 1.5932650057948194e-05
step: 360, loss: 0.16465403139591217
step: 370, loss: 0.00026422415976412594
step: 380, loss: 0.00017880393716041
step: 390, loss: 0.00025748711777850986
step: 400, loss: 0.0002837704960256815
step: 410, loss: 6.83379330439493e-05
step: 420, loss: 0.0012675740290433168
step: 430, loss: 4.947320121573284e-05
step: 440, loss: 0.0001539126824354753
step: 450, loss: 0.2397109717130661
step: 460, loss: 0.050735924392938614
step: 470, loss: 0.004090490750968456
step: 480, loss: 6.297365325735882e-05
step: 490, loss: 0.00036912772338837385
step: 500, loss: 0.001392032252624631
step: 510, loss: 0.0001809524546843022
step: 520, loss: 0.0004521885421127081
step: 530, loss: 0.00010989701695507392
/usr/local/lib/python3.7/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=0.950943396226415, f1=0.9384687646782527,
best_f1=0.9444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006767346058040857
step: 10, loss: 0.0036711962893605232
step: 20, loss: 0.0009905119659379125
step: 30, loss: 0.001599328825250268
step: 40, loss: 0.0013924466911703348
step: 50, loss: 0.0006745065911673009
step: 60, loss: 0.0007601968827657402
step: 70, loss: 0.0002985617029480636
step: 80, loss: 0.00011465429270174354
step: 90, loss: 2.5602268578950316e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.000838143692817539
step: 110, loss: 0.0010234435321763158
step: 120, loss: 4.1940227674786e-05

```

```

step: 130, loss: 0.0003602292272262275
step: 140, loss: 0.04622070491313934
step: 150, loss: 0.000559343839995563
step: 160, loss: 5.9897305618505925e-05
step: 170, loss: 0.00015161816554609686
step: 180, loss: 8.605478797107935e-05
step: 190, loss: 0.0001649943587835878
step: 200, loss: 0.0001793291448848322
step: 210, loss: 0.0001787720393622294
step: 220, loss: 0.00034597795456647873
step: 230, loss: 8.58730127220042e-05
step: 240, loss: 0.0009446758776903152
step: 250, loss: 0.0023282822221517563
step: 260, loss: 0.0016065583331510425
step: 270, loss: 0.034661535173654556
step: 280, loss: 0.0013543760869652033
step: 290, loss: 0.0003477366699371487
step: 300, loss: 0.00038565241266041994
step: 310, loss: 0.011614833027124405
step: 320, loss: 0.00024471955839544535
step: 330, loss: 0.0007168235024437308
step: 340, loss: 0.0008755055605433881
step: 350, loss: 0.00028111462597735226
step: 360, loss: 8.026081195566803e-05
step: 370, loss: 0.00026464497204869986
step: 380, loss: 0.02519017830491066
step: 390, loss: 0.0007336859125643969
step: 400, loss: 0.002555044600740075
step: 410, loss: 3.6300985811976716e-05
step: 420, loss: 0.00011496153456391767
step: 430, loss: 0.0006671060691587627
step: 440, loss: 0.00025964638916775584
step: 450, loss: 4.813914347323589e-05
step: 460, loss: 0.0004107909044250846
step: 470, loss: 1.1957946298934985e-05
step: 480, loss: 0.0004260833957232535
step: 490, loss: 0.00042598502477630973
step: 500, loss: 0.007846334017813206
step: 510, loss: 0.002401969628408551
step: 520, loss: 0.00023308073286898434
step: 530, loss: 0.0021612723357975483
/usr/local/lib/python3.7/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=0.9521597770552717, f1=0.9425925925925925,
best_f1=0.9444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.370756530785002e-05
step: 10, loss: 0.00048340845387429
step: 20, loss: 0.0014788978733122349
step: 30, loss: 0.00030144694028422236
step: 40, loss: 4.6220855438150465e-05
step: 50, loss: 0.00046139484038576484
step: 60, loss: 0.00013359445438254625
step: 70, loss: 3.7283596611814573e-05
step: 80, loss: 0.00025278606335632503
step: 90, loss: 8.652130782138556e-05
step: 100, loss: 0.00014421755622606725
step: 110, loss: 0.0003426856710575521
step: 120, loss: 0.00020936956570949405
step: 130, loss: 0.00909795518964529
step: 140, loss: 5.191103628021665e-05
step: 150, loss: 0.001079176552593708
step: 160, loss: 8.266639633802697e-05
step: 170, loss: 3.0454555599135347e-05
step: 180, loss: 0.0018350580940023065
step: 190, loss: 0.0014902190305292606
step: 200, loss: 0.026424376294016838
step: 210, loss: 0.012111296877264977
step: 220, loss: 0.0008345473324880004
step: 230, loss: 0.00010579688387224451
step: 240, loss: 0.0005329950363375247
step: 250, loss: 7.464743248419836e-05
step: 260, loss: 2.324713750567753e-05
step: 270, loss: 3.219397331122309e-05
step: 280, loss: 3.5760222090175375e-05
step: 290, loss: 0.000836177496239543
step: 300, loss: 0.005184963811188936
step: 310, loss: 0.015639694407582283
step: 320, loss: 0.00022996218467596918
step: 330, loss: 6.195496098371223e-05
step: 340, loss: 3.12152951664757e-05
step: 350, loss: 0.0037667674478143454
step: 360, loss: 0.0003154030709993094
step: 370, loss: 0.0014340004418045282
step: 380, loss: 0.0009246104746125638
step: 390, loss: 1.7817341358750127e-05
step: 400, loss: 0.014460649341344833
step: 410, loss: 4.8233207053272054e-05
step: 420, loss: 0.00015975507267285138
step: 430, loss: 2.0059298549313098e-05
step: 440, loss: 0.0004395870491862297
step: 450, loss: 0.0016460237093269825
step: 460, loss: 0.00048098101979121566
```

```

step: 470, loss: 1.0676578312995844e-05
step: 480, loss: 0.0018032500520348549
step: 490, loss: 0.006646408699452877
step: 500, loss: 0.00021377291704993695
step: 510, loss: 0.00493669556453824
step: 520, loss: 0.0003357178356964141
step: 530, loss: 0.0004996138741262257
/usr/local/lib/python3.7/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=0.9524697110904008, f1=0.9422180801491147,
best_f1=0.9444444444444445
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 [00:33, 169.65it/s]
load_f1 = 0.9523809523809523
real_f1 = 0.9503940658321743
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:30, 144.36it/s]